# Structural and biophysical characterisation of *Mycobacterium tuberculosis* dethiobiotin synthetase

A thesis to fulfil the requirements for the degree of doctor of philosophy

Andrew Thompson

Department of Molecular and Biomedical Science

School of Biological Sciences, The University of Adelaide

2018

# Table of contents

# Abstract

Tuberculosis (TB) remains one of the deadliest infectious diseases worldwide. Although considered treatable and curable, control of this disease is threatened by an increase in antibiotic resistant strains of the causative pathogen, *Mycobacterium tuberculosis (Mtb)*. One method to combat the rise in drug resistant TB is the development of therapeutics that act against novel pathways, such as biotin biosynthesis. Biotin (vitamin B7) is an essential vitamin for all forms of life and is utilized by biotin-dependent enzymes in important cellular processes such as energy metabolism and cell wall maintenance. While most organisms acquire biotin through a combination of transport and biosynthesis, *Mtb* relies solely on biosynthesis for biotin acquisition. The biotin biosynthesis pathway has been determined to be essential to both acute and latent stages of *Mtb* infection, signifying development of inhibitors against this pathway as a potential means for new anti-TB drug discovery efforts. The work in this thesis focuses on dethiobiotin synthetase (*Mt*DTBS), the enzyme that catalyzes the penultimate step of biotin biosynthesis: the conversion of 7,8-diaminopelargonic acid (DAPA) to dethiobiotin (DTB). Here, structural and biophysical investigation of *Mt*DTBS ligand binding gleaned insight into its catalytic mechanism and guided the identification and development of alternate compounds that bind to the enzyme. During the structural characterisation of *Mt*DTBS, an X-ray crystallographic technique was developed to aid the generation of ligand-complexed crystal structures.

Preliminary attempts to analyze the ligand-binding activity of *Mt*DTBS was prevented by the presence of crystallographic precipitants in the active site of the enzyme. Present at high concentrations in the crystallisation methodology, precipitants such as sulfate and citrate bound to the phosphate binding loop (P-loop) of *Mt*DTBS, outcompeting ligands of interest from binding. Chapter 2 contains a manuscript published in *Acta Crystallographica Section D*, wherein a technique termed Precipitant Ligand Exchange was developed to overcome this issue. Here, pre-grown crystals of *Mt*DTBS were taken out of the crystal growth liquor and transferred into a solution that contained the ligand of interest in place of the precipitant. A complex of ADP-*Mt*DTBS was then solved, revealing an unexpected binding mode. Subsequently, multiple X-ray crystallographic models of *Mt*DTBS in complex with a number of ligands were solved and are presented throughout this thesis. Given that crystallographic precipitants are commonly observed in structures deposited in the PDB, this problem is likely to be encountered frequently. Consequently, it is expected that this technique will be of use to researchers throughout the world.

*Mt*DTBS is unique in that it can utilize multiple nucleoside triphosphates (NTPs) to catalyze its reaction. Described as promiscuous binding, this ability was previously identified using biochemical means, and X-ray crystallography was performed with the substrate cytidine triphosphate (CTP). Despite this, the structural and biophysical mechanisms of *Mt*DTBS promiscuity were unknown. Here, initial attempts at observing the promiscuous NTP binding characteristics of *Mt*DTBS were inhibited by the presence of the aforementioned crystallographic precipitants. However, the use of Precipitant Ligand Exchange allowed the generation of investigative structural data. Chapter 3 contains a manuscript published in *ACS Catalysis* which details the use of X-ray crystallography and surface plasmon resonance (SPR) to determine that *Mt*DTBS alternates between two NTP binding modes. CTP binds in a high affinity complex that exhibits slow dissociation as observed by SPR. Electron density pertaining to a single, stable CTP conformation involving specific hydrogen bonding to the enzyme was observed by X-ray crystallography. By contrast, other tested NTPs exhibited fast

association and dissociation kinetics, and much lower affinity. Electron density was only present for the triphosphates of these NTPs, while that pertaining to the nucleoside portion was absent. This lead to the hypothesis that promiscuity occurred primarily through high affinity binding solely between the triphosphates and the *Mt*DTBS P-loop. This discovery highlighted the utility of Precipitant-Ligand Exchange, and raises the possibility that observation of this kind of promiscuity in other enzymes was previously inhibited by crystallographic precipitants. Promiscuity has not been demonstrated with DTBS from other species, however structural comparison revealed several similarities and differences. This encourages further exploration in order to determine the exact structural features required for promiscuity, and its occurrence throughout the wider proteome.

Chapter 4 pertains to the identification and development of inhibitors against *Mt*DTBS. Continuing on from a preliminary *in silico* screen of nucleoside analogues and fragments against *Mt*DTBS, X-ray crystallography screening was used to investigate and confirm the binding mode of hits. Two alternate nucleoside scaffolds were identified to bind to *Mt*DTBS, and it was determined that 2′-deoxycytidine triphosphate exhibited higher affinity for *Mt*DTBS than CTP. While this discovery had implications for the *in vitro* use of energy sources during *Mt*DTBS catalysis, this constituted an improved probe for lead compound development. Additionally, a hit from a high concentration soak with a fragment yielded electron density within the DAPA binding pocket of *Mt*DTBS. Interestingly, this compound was confirmed to be a degradation product of another fragment by NMR, LC/MS and small molecule X-ray crystallography. Chemical modification of this compound was guided by structural biology and led to a large increase in binding potency, however inhibition of the bacterium was not observed. It was hypothesized that both enzyme inhibition and bacterial cell entry should be simultaneously optimised in the future. As several compounds were identified to bind in adjacent positions in the enzyme active site, the possibility of generating a large, high affinity compound by linking smaller compounds together was investigated. Crystal structures with preliminary linked compounds indicated that the linking strategy was not optimised, providing future directions for the project.

Altogether, this thesis represents a detailed investigation of the biophysical and structural characteristics of *Mt*DTBS, with future prospects for the development of inhibitors against this promising anti-TB drug target. Having optimised screening techniques such as those for X-ray crystallography and SPR, this project is well suited for future fragment-based lead discovery endeavours. When considered in conjunction with the additional characterisation of the *Mt*DTBS active site, the identification of new fragment scaffolds would significantly broaden the scope for inhibitor development. With the capability for rapid and accurate testing of inhibitors against both the *Mt*DTBS enzyme and *Mycobacterium tuberculosis*, the results obtained from this project provide an excellent opportunity to pursue the development of anti-TB compounds.

# Public Research Output

Listed here is the public research output completed during this Doctor of Philosophy degree. First are the publications presented in this thesis, followed by X-ray crystallographic structures submitted to the Protein Data Bank (PDB) as well as oral and poster presentations.

## Publications

Thompson, A. P., Salaemae, W., Pederick, J. L., Abell, A. D., Booker, G. W., Bruning, J. B., Wegener, K. L., Polyak, S. W. (2018). Mycobacterium tuberculosis Dethiobiotin Synthetase Facilitates Nucleoside Triphosphate Promiscuity through Alternate Binding Modes. *ACS Catalysis*, *8*(11), 10774–10783.

- Impact factor (2017): 11.384

Thompson, A. P., Wegener, K. L., Booker, G. W., Polyak, S. W., & Bruning, J. B. (2018). Precipitant–ligand exchange technique reveals the ADP binding mode in Mycobacterium tuberculosis dethiobiotin synthetase. *Acta Crystallographica Section D: Structural Biology*, *74*(10), 965–972.

- Impact factor (2017): 3.099

Thompson, A. P., Sternicki, L. M., Wegener, K. L., Wei, L., Zhu, L. Booker, G. W., Polyak, S. W., Li, Y. (2016) Biotin Biology as a Target for New Anti-Tuberculosis Drugs. *Jiangsu J. Prev. Med.*, **27** (3): 257-261.

## PDB Submissions

| **ID** | **Title** |
|---|---|
| 6CZD | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with adenosine diphosphate |
| 6E05 | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with cytidine triphosphate solved by precipitant-ligand exchange (crystals grown in sulfate precipitant) |
| 6E06 | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with cytidine triphosphate solved by precipitant-ligand exchange (crystals grown in citrate precipitant) |
| 6CVE | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin Synthetase in complex with cytidine triphosphate and 7,8-diaminopelargonic acid |
| 6CVU | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with cytidine |
| 6CVF | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with cytidine diphosphate |

| | |
|---|---|
| 6CVV | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with adenosine triphosphate (ATP) - promiscuous binding mode with disordered nucleoside |
| 6CZE | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with inosine triphosphate (ITP) - promiscuous binding mode with disordered nucleoside |
| 6CZB | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with uridine triphosphate (UTP) - promiscuous binding mode with disordered nucleoside |
| 6CZC | Crystal structure of *Mycobacterium tuberculosis* dethiobiotin synthetase in complex with thymidine triphosphate (TTP) - promiscuous binding mode with disordered nucleoside |

## Oral presentations

Structural studies reveal unconventional modes of nucleotide triphosphate binding for *M. tuberculosis* dethiobiotin synthetase: a novel antibiotic target
*Talk presented at European Molecular Biology Laboratory Postgraduate Symposium (Garvan Institute, Sydney, Australia, 2017)*

Structural studies reveal unconventional modes of nucleotide triphosphate binding for *M. tuberculosis* dethiobiotin synthetase: a novel antibiotic target.
*Invited talk presented at Assoc. Prof. Jaime Triccas Laboratory (Charles Perkins Centre, The University of Sydney Sydney, Australia 2017)*

Using chemical probes to map the active site of dethiobiotin synthetase from *M. tuberculosis*.
*Lightning talk presented at ComBio2017 Conference (Adelaide, Australia, 2017)*

Structurally guided insights into the development of novel anti-TB compounds.
*Invited talk presented at Prof. David Baker Laboratory (University of Washington, Seattle, WA, USA, 2016)*

*S*tructurally guided insights into the development of novel anti-TB compounds.
*Invited talked presented at Prof. Sachdev Sidhu Laboratory (The Donnelly Centre, University of Toronto, Toronto, ON, Canada, 2016)*

An *in silico* fragment-based approach to developing new anti-TB drugs.
*Poster teaser presented at Fragment-Based Drug Discovery Down Under Conference (Monash Institute of Pharmaceutical Sciences, Melbourne, 2016)*

## Poster Presentations

**Thompson, A.P.**, Salaemae, W., Gaiser, B., Lee, J.K., Abell, A.D., Booker, G.W., Bruning, J.B., Polyak, S.W., Wegener, K.L. Using chemical probes to map the active site of dethiobiotin synthetase from Mycobacterium tuberculosis
*Presented at Lorne Proteins Conference (Lorne, Australia, 2018)*

**Thompson, A.P.**, Salaemae, W., Gaiser, B., Lee, J.K., Abell, A.D., Booker, G.W., Bruning, J.B., Polyak, S.W., Wegener, K.L. Using chemical probes to map the active site of dethiobiotin synthetase from Mycobacterium tuberculosis

*Presented at ComBio2017 Conference (Adelaide, Australia, 2017 -* ***awarded poster prize****)*

**Thompson, A.P.**, Salaemae, W., Gaiser, B., Lee, J.K., Abell, A.D., Booker, G.W., Bruning, J.B., Polyak, S.W., Wegener, K.L. Using chemical probes to map the active site of dethiobiotin synthetase from Mycobacterium tuberculosis

*Presented at Adelaide Protein Group (Adelaide, Australia, 2017 -* ***awarded poster prize****)*

**Thompson, A. P.**, Salaemae, W., Booker, G. W., Bruning, J. B., Wegener, K. L., Polyak, S. W. Crystallographic confirmation of *in silico* fragment hits generates linking strategies for the development of novel anti-tuberculosis agents

*Presented at EMBL Australia Postgraduate Conference (SAHMRI, Adelaide, 2016)*

**Thompson, A. P.**, Salaemae, W., Booker, G. W., Bruning, J. B., Wegener, K. L., Polyak, S. W. An *in silico* fragment-based approach to developing new anti-TB drugs.

*Presented at Fragment-Based Lead Discovery Conference (Boston, Massachusetts, USA, 2016)*

**Thompson, A. P.**, Salaemae, W., Booker, G. W., Bruning, J. B., Wegener, K. L., Polyak, S. W. An *in silico* fragment-based approach to developing new anti-TB drugs.

*Presented at Fragment-Based Drug Discovery Down Under (Monash Institute of Pharmaceutical Sciences, Melbourne, Australia, 2016)*

**Thompson, A. P.**, Salaemae, W., Booker, G. W., Bruning, J. B., Wegener, K. L., Polyak, S. W. Crystallographic study of *in silico* derived *Mycobacterium tuberculosis* Dethiobiotin Synthetase Inhibitor Fragments.

*Presented at Adelaide Protein Group Conference (Adelaide, Australia, 2016)*

**Thompson, A. P.**, Salaemae, W., Booker, G. W., Bruning, J. B., Wegener, K. L., Polyak, S. W. Crystallographic study of *in silico* derived *Mycobacterium tuberculosis* Dethiobiotin Synthetase Inhibitor Fragments.

*Presented at Lorne Conference on Protein Structure and Function (Lorne, Australia, 2016)*

Salaemae, W., **Thompson, A. P.**, Booker, G. W., Wegener, K. L., Polyak, S. W. Investigation into the Mechanism of Mycobacterium tuberculosis Dethiobiotin Synthetase Inhibitor Fragments Using NMR Based Techniques.

*Presented at Australian Society of Medical Research Conference (Adelaide, Australia, 2016)*

Salaemae, W., **Thompson, A. P.**, Booker, G. W., Wegener, K. L., Polyak, S. W. Investigation into the Mechanism of Mycobacterium tuberculosis Dethiobiotin Synthetase Inhibitor Fragments Using NMR Based Techniques.

*Presented at ComBio (Melbourne, Australia, 2015)*

Salaemae, W., **Thompson, A. P.**, Booker, G. W., Wegener, K. L., Polyak, S. W. Investigation into the Mechanism of Mycobacterium tuberculosis Dethiobiotin Synthetase Inhibitor Fragments Using NMR Based Techniques.
*Presented at International Anti-Microbial Resistance Workshop (Adelaide, Australia, 2015)*

Salaemae, W., **Thompson, A. P.**, Booker, G. W., Wegener, K. L., Polyak, S. W. Investigation into the Mechanism of Mycobacterium tuberculosis Dethiobiotin Synthetase Inhibitor Fragments Using NMR Based Techniques.
*Presented at Melbourne Protein Group (Melbourne, Australia 2015)*

Salaemae, W., **Thompson, A. P.**, Booker, G. W., Wegener, K. L., Polyak, S. W. Investigation into the Mechanism of Mycobacterium tuberculosis Dethiobiotin Synthetase Inhibitor Fragments Using NMR Based Techniques.
*Presented at Adelaide Protein Group (Adelaide, Australia, 2015)*

# Statement of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time. I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

10/12/2018

**Andrew Thompson**

# Acknowledgements

Thank you to Grant, Kate, John and Steven. Your guidance was invaluable throughout my PhD, especially in writing this thesis and the papers within it. Again, thank you.

Cheers to all past and present members of the Booker and Bruning labs, and to those from other labs around the MLS.

Thanks to my friends and family.

Thank you, Eleni.

Above all, thanks Mum and Dad.

# List of Abbreviations

| | |
|---|---|
| ADE | Adenosine |
| ADP | Adenosine diphosphate |
| ATP | Adenosine triphosphate |
| BME | β-mercaptoethanol |
| °C | Degrees Celsius |
| CTN | Cytidine |
| CDP | Cytidine diphosphate |
| CTP | Cytidine triphosphate |
| DAPA | 7,8-diaminopelargonic acid |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DTB | Dethiobiotin |
| DTBS | Dethiobiotin synthetase |
| DTT | Dithiothreitol |
| EDTA | Ethylene diamine tetra-acetic acid |
| FBLD | Fragment-based lead discovery |
| g | Gram |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HTS | High throughput screening |
| KCl | Potassium chloride |
| kDa | Kilo Dalton |
| $K_D$ | Affinity constant |
| $k_d$ | Kinetic dissociation constant |
| $k_a$ | Kinetic association constant |
| $K_i$ | Inhibition constant |
| LB | Luria broth |
| m | milli- |
| M | Molar |
| Mg | Magnesium |
| $MgCl_2$ | Magnesium chloride |
| min | Minute |
| *Mt* | *Mycobacterium tuberculosis* |
| MW | Molecular weight |
| n | nano- |
| N- | amino- |
| NTP | Nucleoside triphosphate |
| NpP | Nucleoside polyphosphate |
| PMSF | Phenylmethylsulfonylfluoride |
| $PO_4^{2-}$ | Phosphate ion |
| Rpm | Rotations per minute |
| SEM | Standard error of the mean |
| $SO_4^{2-}$ | Sulfate ion |
| SPR | Surface plasmon resonance |
| TB | Tuberculosis |
| TBS | Tris buffered saline |
| Tris | 2-amino-2-hydroxymethylpropane-1,3-diol |
| µ | Micro- |
| Vmax | Maximum velocity |

# Chapter 1:

## Introduction

## 1.1 Thesis description

This thesis details the development and use of X-ray crystallography and surface plasmon resonance protocols to aid in the development of an antibiotic lead molecule. These techniques were used to investigate the structure and function of the anti-tuberculosis target enzyme: *Mycobacterium tuberculosis* dethiobiotin synthetase (*Mt*DTBS). The insights gained here were then applied to lead molecule development, resulting in promising candidates for future antibiotic development. A central theme of this work is that the fragment-based lead discovery process is best guided by structural and functional information determined by a combination of orthogonal techniques.

This thesis is presented as a combination of published manuscripts and traditional thesis chapters. Due to this format, each results chapter contains its own materials and methods section. Similarly, references for each manuscript are included within those chapters (1, 2 and 3) and any remaining references (from chapters 1, 4 and 5) appear at the end of the thesis.

The introduction chapter discusses multiple aspects of antibiotic drug development. Included is an invited review published in the *Jiangsu Journal of Preventative Medicine* as a result of collaboration with Yan Li, a visiting scholar in the Booker Laboratory. This review describes the global healthcare impact of tuberculosis and introduces biotin biology as a pathway for anti-mycobacterial drug development. Following this review, more specific background is included regarding *Mt*DTBS. The chapter concludes with discussion of the methods of lead molecule identification and development, with a specific focus on fragment-based lead discovery (FBLD).

Chapter 2 is a manuscript entitled "Precipitant-ligand exchange technique reveals ADP binding mode in *Mt*DTBS*"* that has been published in *Acta Crystallographica Section D*. This manuscript details the development of a precipitant-ligand exchange crystallographic soaking technique. This technique allowed the generation of atomic resolution structures of proteins in complex with ligands that would otherwise be occluded by crystallographic reagents such as precipitants. The technique was validated by replicating a previous structure of *Mt*DTBS with cytidine triphosphate. Subsequently, a novel binding site for adenosine diphosphate, an inhibitor and reaction by-product of *Mt*DTBS was discovered. As crystallographic precipitants are commonly present in structures found in the Protein Data Bank, it is anticipated that this methodology will be useful to the wider crystallographic community.

Chapter 3 is a manuscript entitled "*Mycobacterium tuberculosis* dethiobiotin synthetase facilitates nucleoside triphosphate promiscuity through alternate binding modes" that has been published in the American Chemical Society Journal *Catalysis*. In this work, precipitant-ligand exchange, as described in Chapter 2, was used to generate atomic resolution data that revealed the mechanism of promiscuous nucleoside triphosphate (NTP) utilization by *Mt*DTBS. This phenomenon required careful analysis of crystallographic structural data and the accompanying biophysical data generated by surface plasmon resonance. Not only does this paper contribute to the overall understanding of substrate utilization by enzymes, but it also examines the *Mt*DTBS binding site in great detail, which aided lead molecule design in the following chapter.

Chapter 4 pertains to the identification of compounds that bind to *Mt*DTBS, followed by their structure guided chemical modification towards high affinity binding. X-ray crystallography

screening was used to identify an optimised nucleoside ligand and a novel small molecule that bound to the active site. Interestingly, this small molecule was confirmed to be a degradation product of another fragment by NMR, LC/MS and small molecule X-ray crystallography. Chemical modification guided by structural biology led to a large increase in binding affinity. As several compounds were identified to bind in adjacent positions in the enzyme active site, the possibility of generating a high affinity compound by linking smaller compounds together was investigated. Crystal structures with preliminary linked molecules revealed strategies for future development.

Chapter 5 features an overall discussion of the work presented in this thesis, as well as the future directions of this project. Advancements in the characterisation of *Mt*DTBS are discussed in the context of the broader literature. This is then followed by an analysis of how this information guided and will continue to guide the development of the described compound series towards becoming inhibitors of *Mt*DTBS and *Mycobacterium tuberculosis*.

## 1.2 Statement of authorship

| | |
|---|---|
| Title of Paper | Biotin Biology as a Target for New Anti-Tuberculosis Drugs |
| Publication Status | Published |
| Publication Details | Thompson, A. P., Sternicki, L. M., Wegener, K. L., Wei, L. U., Limei, Z. U. O., Booker, G. W., Polyak, S. W. & Yan, L. I. (2016). *Jiangsu Journal of Preventive Medicine*. 27, 257–261. |

### Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Andrew P. Thompson | | |
| Contribution to the Paper | Prepared manuscript and figures. | | |
| Overall percentage (%) | 80% | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature | | Date | 20/11/18 |

### Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| | | | |
|---|---|---|---|
| Name of Co-Author | Louise M. Sternicki | | |
| Contribution to the Paper | Manuscript and figure preparation regarding biotin protein ligase. | | |
| Signature | | Date | 20/11/18 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Kate L. Wegener | | |
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 5/12/18 |

| Name of Co-Author | L. U. Wei | | |
|---|---|---|---|
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 5/12/2018 |

| Name of Co-Author | Z. U. O. Limei | | |
|---|---|---|---|
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 5/12/2018 |

| Name of Co-Author | Grant W. Booker | | |
|---|---|---|---|
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 20/11/2018 |

| Name of Co-Author | Steven W. Polyak | | |
|---|---|---|---|
| Contribution to the Paper | Conceived *Mt*DTBS project. Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 23/11/2018 |

| Name of Co-Author | Yan Li | | |
|---|---|---|---|
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 5/12/2018 |

·结核病专题论著·

# 生物素——抗结核药物新研究方向

THOMPSON Andrew P[1], STERNICKI Louise M[1], WEGENER Kate L[1],
陆伟[2], 竺丽梅[2], BOOKER Grant W[1], POLYAK Steven W[1], 李燕[2]
1. 澳大利亚阿德莱德大学生命科学院, 阿德莱德 5005; 2. 江苏省疾病预防控制中心, 南京 210009

**摘要**: 耐药菌的出现对全球公共卫生构成了巨大的挑战, 结核分枝杆菌(TB)是结核病的病原菌, 是临床上最重要的病原体之一。新抗生素的研发用以治疗耐药性结核病已成为临床及公共卫生管理亟待解决的重要问题。生物素(维生素 B7)在结核杆菌的代谢途径如脂肪酸合成和三羧酸循环中发挥重要作用, 通过生物素蛋白连接酶连接活性辅助因子而被激活。生物素蛋白连接酶及其相关生物合成通路是微生物合成生物素过程中必不可或缺的关键分子, 因而有可能成为新的药物靶点。遗传研究表明, 结核杆菌需要由菌体进行生物素的全程合成, 因此, 生物素合成酶在在结核杆菌代谢过程中起着至关重要的作用, 引起了抗结核药物研究者的高度重视。本文论述了结核杆菌生物素合成通路, 并重点分析了生物素合成过程中可能成为药物靶点的关键酶, 旨在为结核病的治疗提供新思路, 具有重要的临床意义。

**关键词**: 结核分枝杆菌; 耐药菌; 生物素; 药物靶点
**中图分类号**: R52　　**文献标识码**: A　　**文章编号**: 1006－9070(2016)03－0257－05

## Biotin Biology as a Target for New Anti-Tuberculosis Drugs

THOMPSON Andrew P*, STERNICKI Louise M, WEGENER Kate L, LU Wei,
ZHU Li-mei, BOOKER Grant W, POLYAK Steven W, LI Yan
* *University of Adelaide, School of Biological Sciences, Adelaide* 5005, *Australia*

**Abstract**: The rise of drug resistant bacteria poses a great global healthcare challenge. One of the most clinically important pathogens is *Mycobacterium tuberculosis*, the agent responsible for tuberculosis (TB). New classes of antibiotics with novel mechanisms of action are required to combat drug resistant TB. The biotin biocycle presents multiple targets for antimicrobial drug discovery, namely the biotin-dependent enzymes biotin protein ligase and the biotin biosynthesis pathway. Biotin (vitamin B7) is a cofactor for enzymes that play key roles in important metabolic pathways in TB, such as fatty acid synthesis and replenishing the citric acid cycle. The biotin dependent-enzymes are activated via the attachment of the cofactor through the activity of biotin protein ligase. Genetic studies suggest that *M. tuberculosis* fulfils its biotin requirements solely through *de novo* synthesis, making the biotin synthetic enzymes attractive for drug discovery. Here we review the biotin biocycle and highlight studies investigating the key enzymes as possible drug targets. This knowledge underpins research efforts to exploit these in the development of new treatments for TB.

**Key words**: Mycobacterium tuberculosis; Drug resistant bacteria; Biotin; Drug targets

### 1　Tuberculosis and Drug Resistance

Tuberculosis (TB) is an infectious disease primarily caused by the pathogenic bacterium *Mycobacterium tuberculosis*. TB is a major global health care issue, with an estimated one third of the world's population infected by the latent organism[1]. The risk of becoming ill with TB after bacterial infection is only 10%, but because the number of people infected with the pathogen is so large, TB is ranked alongside HIV/AIDS as a leading cause of deaths worldwide. In 2014, 9.6 million people contracted TB, resulting in 1.5 million deaths[1]. *M. tuberculosis* colonizes the lungs of infected patients, and is transmitted via droplets in the air when infected people cough, sneeze or spit. It is a highly resilient bacterium capable of inhabiting cells in the lung[2] or macrophages where it can survive in a dormant state for extended periods of time[3] and subsequently reactivate to cause TB[2]. Its unique cell wall structure, consisting primarily of mycolic acids, is thought to contribute to its pathogenicity[4-5].

Treatment for TB is a six-month drug regimen-requiring front line antibiotics, such as isoniazid and rifampicin[1]. The extended length of treatment makes patient compliance a major issue, especially if appropriate guidance and support are not available. Poor patient adherence and inappropriate use of the drugs are cited as major factors for the development

DOI: 10.13668/j.issn.1006－9070.2016.03.01
**基金项目**: National Health and Medical Research Council of Australia (APP1068885); National Natural Science Foundation of China (No. 81301448); Key Medical Talent Foundation of Jiangsu Provincial Center for Disease Control and Prevention (No. JKRC20110029); Jiangsu health international exchange program sponsorship (JSH－2015－002).
**作者简介**: Thompson Andrew P(1991－), 男, 澳大利亚人, 博士在读, 主要从事以生物素合成酶为靶点的新型药物筛选工作。
**通讯作者**: Polyak Steven W, E-mail: steven.polyak@adelaide.edu.au; 李燕, E-mail: liyan.nju@163.com

of multidrug resistant TB (MDR-TB). MDR-TB is caused by *M. tuberculosis* strains that are unresponsive to isoniazid and rifampicin. This form of TB requires second line treatments that often have greater associated toxicity. *M. tuberculosis* strains that are resistant to a larger selection of anti-TB drugs, including second line treatments, have recently been documented; these have been termed extensive drug resistant TB (XDR-TB). Approximately 480 000 people contracted MDR-TB in 2014, with an estimated 9.7% of those cases being XDR-TB.

## 2 Biotin biology and therapeutic intervention

The rise of drug resistant *M. tuberculosis* highlights a need for safer and more effective treatment options. Biotin biosynthesis and utilisation pathways provide an array of potential drug targets for new antimicrobial therapies that could fulfil this need. Biotin is an essential vitamin that is utilised by all living organisms. It is a well-studied enzyme cofactor that is required for certain carboxylation, decarboxylation and transcarboxylation reactions. *M. tuberculosis* possesses two classes of biotin-dependent enzymes that play key metabolic roles in fatty acid metabolism and energy production, namely Acyl-CoA carboxylases (ACCs) and pyruvate carboxylases (PC). These enzymes are activated when ligated to biotin by the enzyme biotin protein ligase (BPL). Here we discuss biotin-dependent enzymes, BPL and biosynthetic enzymesas possible drug targets for the treatment of TB.

2.1 Biotin-dependent enzymes ACCs are a class of enzyme that catalyse the carboxylation of short chain acyl-CoA substrates[6]. The products of these reactions are used in the synthesis of fatty acids and polyketides, including the pathogenically important mycolic acids[4-5]. Mycolic acids are unique to mycolata species, and are a major component of the complex cell wall of *M. tuberculosis*. Multiple anti-TB drugs currently on the market work by inhibiting lipid synthesis, such as isoniazid, ethionamide and thiocarlide, demonstrating the efficacy of such an approach[7]. Isoniazid and ethionamide target an enoyl-acyl carrier protein reductase known as InhA, preventing the synthesis of fatty acids including mycolic acid[8]. Similarly, thiocarlide inhibits the production of oleic acid[9]. The ACCs as targets for new antibiotics have recently been reviewed[10].

PC catalyses the conversion of pyruvate to oxaloacetate, an intermediate of the citric acid cycle[7], and therefore plays a pivotal role in the energy metabolism of cells. *M. tuberculosis* has three oxaloacetate producing enzymes, and the efficacy of inhibiting PC for the treatment of TB has not yet been confirmed[7].

2.2 Biotin protein ligase (BPL) All organisms contain between one and five biotin-dependent enzymes, which each require BPL for biotin ligation[11-12]. Biotin is attached to a specific lysine residue in the active site of the biotin-dependent enzymes. Biotinylation proceeds via a two-step reaction, requiring both biotin and ATP. These reactants first bind to their adjacent binding pockets, yielding the reaction intermediate biotinyl-5′-AMP. The BPL: biotinyl-5′-AMP complex then interacts with the biotin-dependent enzyme such that the target lysine side chain enters the BPL active site. The positively charged ε-amino group of the lysine is then selectively biotinylated[12-13]. Biotin-dependent enzymes require biotinylation for their activity, and without it they are unable to fulfill their important metabolic roles. Therefore, small molecule inhibitors that target BPL represent possible treatments for TB.

Validation of BPL as a drug target has been provided by both genetic and pharmacological studies. Knockout studies showed that deletion of the BPL gene in fast growing model species *Mycobacterium smegmatis*, prevented colony growth, demonstrating that BPL is essential for viability *in vitro*[14]. Two non-hydrolysable analogues of the biotinyl-5′-AMP reaction intermediate have also been shown to pharmacologically inhibit bacterial growth[14-15]. Importantly, these compounds are also potent antibacterials against various drug-resistant *M. tuberculosis* strains including MDR-TB and XDR-TB[14]. Together these studies identify *M. tuberculosis* BPL as an attractive target for antibiotic development.

Certain bacterial BPLs, including the *M. tuberculosis* enzyme, possess an ordered binding mechanism. This knowledge impacts upon the design of inhibitors that target BPL. Here biotin binds to the enzyme first, causing a disordered to ordered conformational change within the enzyme. Restructuring of a biotin-binding loop, positions the side chain of a key tryptophan residue such that it is conducive for ATP binding (figure 1). Binding of both biotin and ATP is required for the synthesis of biotinyl-5′-AMP. Restructuring of a second disordered adenylate-binding loop further stabilizes the interaction with the reaction intermediate[12,16-19]. As the ATP pocket is unstructured without biotin, it cannot be targeted on its own. Therefore, efforts to design BPL inhibitors that target the active site must contain a biotinyl functional group, and this serves as a starting point for the design of new compounds[15-16,20-21].

2.3 Biotin biosynthesis Organisms acquire biotin via a variety of different mechanisms. In humans and other mammals, biotin is obtained through exogenous

sources such as diet or from biotin-producing bacteria in the gut[22-23]. Certain organisms have the metabolic pathways required for the synthesis of biotin, including some plants, fungi, bacteria and other microorganisms[24]. Bacteria can also scavenge extracellular biotin through biotin ATP-dependent transporters. These include BioY, YigM and MadN, which have been characterised in various bacteria[25]. Genome annotation studies and nucleotide sequence analysis with homology algorithms have failed to identify BioY, YigM or MadN homologues in *M. tuberculosis*[7,26-27]. These findings suggest that *M. tuberculosis* relies on *de novo* synthesis as its sole source of biotin, highlighting this pathway as a potential drug target.

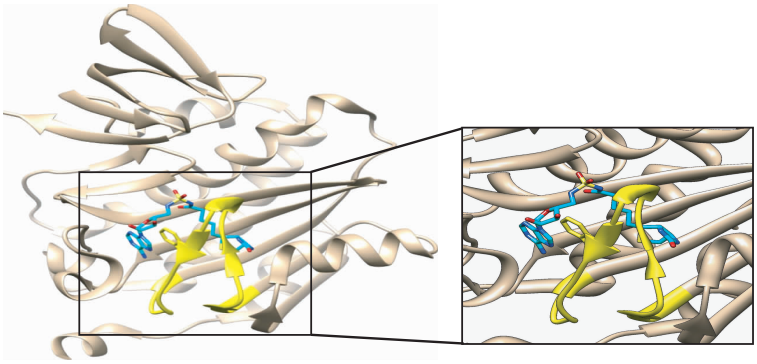


Figure 1　Structure of *M. tuberculosis* BPL with a bisubstrate inhibitor (cyan) that mimics the reaction intermediate (PDB:3RUX)[14]. This demonstrates adjacent binding pockets for biotin (right) and ATP (left). The biotin-binding loop (yellow) becomes structured after biotin binding, allowing the ATP pocket to form and repositioning a Tryptophan residue (side chain shown) that is necessary for ATP binding. This mechanism results in the ordered binding of biotin prior to ATP. For details in picture printed in colors, please see back cover.

The biotin biosynthesis pathway is responsible for the enzymatic conversion of a malonyl CoA precursor to biotin (figure 2). Briefly, the biotin biosynthetic pathway begins with *O*-methyltransferase (encoded by *bioC*) converting malonyl CoA to malonyl CoA methyl ester, followed by elongation of the alkyl chain using the fatty acid synthesis pathway. Pimeloyl-ACP methyl ester is exited from the fatty acid synthesis pathway and carboxylesterase (encoded by *bioH*) converts it to pimeloyl-ACP. The next four reactions are conserved between all biotin producing organisms. 7-keto-8-aminopelargonic acid (KAPA) is produced by KAPA synthase (encoded by *bioF*) through the use of L-alanine. KAPA is subsequently converted to 7-8-diaminopelargonic acid (DAPA) by DAPA synthase (encoded by *bioA*), utilising S-adenyl methionine (SAM) as an amino donor. Dethiobiotin synthetase (DTBS, encoded by *bioD*) then utilises $CO_2$ to close the ureido ring of DAPA in an ATP dependent manner to form dethiobiotin[28]. Finally, a sulphur and two electrons from flavodoxin, SAM and nicotinamide adenine dinucleotide phosphate (NADPH) are utilised by biotin synthase to catalyse the closure of the thiophane ring, resulting in the formation of d-biotin. For a more detailed explanation of this metabolic pathway we encourage readers to see our recent review[7].

Several studies provide strong genetic and biochemical evidence for the inhibition of biotin biosynthesis as a strategy for developing anti-TB agents. Genetic knockout studies using strains of *M. tuberculosis* deficient in the *bioA* and *bioF* genes are unable to colonisemice in models of infection[29-30]. A conditional *bioA* knockout reveals that this gene is also essential for persistence as a chronic infection[29]. Similarly, genetic studies in *M. smegmatison* the biotin biosynthesis genes *bioA*, *bioF* and *bioB* confirm that the biotin biosynthesis pathway is essential for *in vitro* growth and *in vivo* pathology[31-33]. Finally, knockout of a reductase required for the synthesis of the biotin precursor pimeloyl-thioester, reduced growth of *Mycobacterium marinum* on blood agar, unless supplemented with high levels of biotin[34]. These studies indicate biotin biosynthesis is a promising drug target for TB infection.

Biotin biosynthesis offers many attractive features as an anti-TB drug target. Foremost of these is the lack of a biotin biosynthetic pathway in mammals. The absence of human homologues means that it is less likely that the drug will have unintended side effects, as commonly seen with current anti-TB drugs. In addition, *M. tuberculosis* is unable to scavenge biotin from the surrounding environment and is solely dependant on *de novo* synthesis. Finally, the biotin synthesis pathway appears to be essential for *M. tuberculosis* survival in the dormancy state. Agents capable of targeting the latent bacterial stage are highly desired as most antibiotics are ineffective against this stage of its lifecycle. Treatments that are active against this stage would allow infections to be resolved before becoming symptomatic.

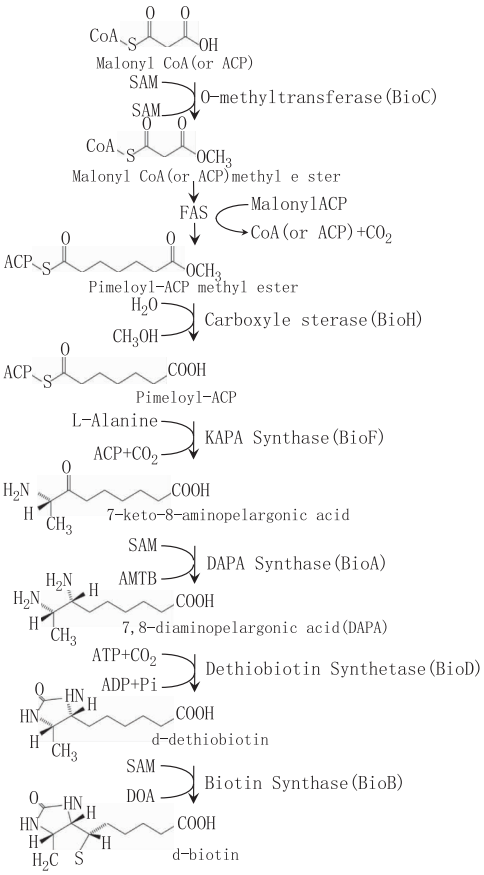



Figure **2** Schematic of the substrates and products of the biotin biosynthesis pathway, with the catalysing enzyme and its encoding gene indicated. Adapted from Salaemae et al[7]

## 3 Antibiotic Discovery Approaches

The rise of antibiotic resistance constitutes a major health concern, especially considering the lack of antibiotic drug development in recent years. Front line drugs with efficacy against drug resistant strains and that are safer than current treatments are desperately required. Treatments that also target dormant *M. tuberculosis* are also needed. Towards this end, we must discover new agents with novel mechanisms of action, that will not subject to existing resistance mechanisms. The biotin biocycle provides several promising new and highly attractive drug targets for TB research. However, translation of our knowledge about this pathway into novel drugs is still challenging. Antibiotic discovery should be viewed as a highly specialised area research where the rules that govern development of drugs for other diseases areas do not necessarily apply. A recent review from Astra Zeneca detailed their extensive antibiotic drug discovery programs from the last decade, from which no successful drugs were produced[25]. This review highlighted many issues with current antibiotic drug discovery strategies, culminating in the conclusion that these approaches are often not well suited to antibiotic discovery. High rates of attrition arose predominantly at the step of converting a molecule with bioactivity against a purified protein into an antibacterial with whole cell activity. This was due, in part, to poor understanding of how compounds interact with microbes, such as the mechanisms of cell wall penetration and efflux[34]. Also, medicinal chemistry approaches to chemical optimisationcan be problematic, including a tendency to engineer compounds with higher *in vitro* potency that coincides with higher lipophilicity. Antibiotics have been shown to occupy a different chemical spacefrom other drugs and do not abide by the rules and metrics that define traditional 'drug-like' molecules[35]. Consequently, compound libraries that have been curated for other disease areas are often not ideal for antibacterial research. The challenge now is to assemble new libraries with the appropriate chemical diversity for antibiotic discovery[36]. Fragment based libraries, for example, can be investigated as these cover large regions of chemical space. The small molecule compounds obtained from these screens provide starting points for medicinal chemistry to develop more drug-like compounds through chemical extension or concatenation. This structure-guided approach is yet to be fully applied to antibiotic discovery. Natural product libraries also have strong potential, with the biological world already providing many important chemical scaffolds with potential as antibacterial drugs, such as the recent discovery of secondary metabolite texiobactin[37]. Drug screening projects using novel and diverse compound libraries, and that contain bioactive compounds that target biotin biosynthesis and its utilisation, promise exciting new avenues for TB drug discovery. Researchers in China and around the world continue to search niche environments for new bioactive compounds, for example in the South China Sea or from plants used in traditional Chinese medicine. Towards this end, Chinese researchers in partnership with international collaborators can have an impact on the world stage. It is through multidisciplinary research that we have the best opportunity to finally combat TB.

**Acknowledgements**: We are grateful to the Wallace and Carthew families for their financial support of this work.

**Reference**

[1] World Health Organisation. Global tuberculosis report[R]. Geneva: WHO, 2015.

[2] SINGH G, SINGH G, JADEJA D, et al. Lipid hydrolizing enzymes in virulence: Mycobacterium tuberculosis as a model system[J]. Crit Rev Microbiol, 2010, 36(3): 259.

[3] PODINOVSKAIA M, LEE W, CALDWELL S, et al. Infection of macrophages with Mycobacterium tuberculosis induces global modifications to phagosomal function: Mtb infection modifies phagosome function[J]. Cell Microbiol, 2013, 15(6): 843.

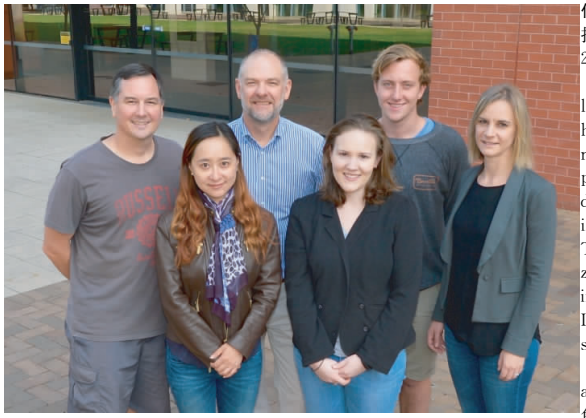
[4] TAKAYAMA K, WANG C, BESRA GS. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis [J]. Clin Microbiol Rev, 2005, 18(1): 81.

[5] BRENNAN PJ, CRICK DC. The cell-wall core of Mycobacterium tuberculosis in the context of drug discovery[J]. Curr Top Med Chem, 2007, 7(5): 475.

[6] SALAEMAE W, AZHAR A, BOOKER GW, et al. Biotin biosyn-

thesis in Mycobacterium tuberculosis: physiology, biochemistry and molecular intervention[J]. Protein Cell, 2011, 2(9): 691.
[7] SALAEMAE W, BOOKER GW, POLYAK SW. The role of biotin in bacterial physiology and virulence: a novel antibiotic target for Mycobacterium tuberculosis[J]. Microbiolspec, 2016, 4: 2.
[8] MARRAKCHI H, LANÉELLE G, QUÉMARD A. InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II[J]. Microbiology, 2000, 146(Pt 2): 289.
[9] PHETSUKSIRI B, JACKSON M, SCHERMAN H, et al. Unique mechanism of action of the thiourea drug isoxyl on Mycobacterium tuberculosis[J]. J Biol Chem, 2003, 278(52): 53123.
[10] POLYAK SW, ABELL AD, WILCE MC, et al. Structure, function and selective inhibition of bacterial acetyl-coa carboxylase [J]. Appl Microbiol Biotechnol, 2012, 93(3): 983.
[11] CHAPMAN-SMITH A, CRONAN JE. Molecular biology of biotin attachment to proteins[J]. J Nutr, 1999, 129(Suppl2): 477.
[12] PAPARELLA AS, SOARES DA COSTA TP, YAP MY, et al. Structure guided design of biotin protein ligase inhibitors for antibiotic discovery[J]. Curr Top Med Chem, 2014, 14(1): 4.
[13] BAGAUTDINOV B, MATSUURA Y, BAGAUTDINOVA S, et al. Protein biotinylation visualized by a complex structure of biotin protein ligase with a substrate[J]. J Biol Chem, 2008, 283 (21): 14739.
[14] DUCKWORTH BP, GEDERS TW, TIWARI D, et al. Bisubstrate adenylation inhibitors of biotin protein ligase from Mycobacterium tuberculosis[J]. Chem Biol, 2011, 18(11): 1432.
[15] TIEU W, POLYAK SW, PAPARELLA AS, et al. Improved synthesis of biotinol-5′-AMP: implications for antibacterial discovery [J]. ACS Med Chem Lett, 2015, 6(2): 216.
[16] SOARES DA COSTA TP, TIEU W, YAP MY, et al. Selective inhibition of biotin protein ligase from Staphylococcus aureus [J]. J Biol Chem, 2012, 287(21): 17823.
[17] MA Q, AKHTER Y, WILMANNS M, et al. Active site conformational changes upon reaction intermediate biotinyl-5′-AMP binding in biotin protein ligase from Mycobacterium tuberculosis [J]. Protein Sci, 2014, 23(7): 932.
[18] WEAVER LH, KWON K, BECKETT D, et al. Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator[J]. Proc Natl Acad Sci U S A, 2001, 98(11): 6045.
[19] WOOD ZA, WEAVER LH, BROWN PH, et al. Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution[J]. J Mol Biol, 2006, 357(2): 509.
[20] SOARES DA COSTA TP, TIEU W, YAP MY, et al. Biotin analogues with antibacterial activity are potent inhibitors of biotin protein ligase[J]. ACS Med Chem Lett, 2012, 3(6): 509.
[21] TIEU W, JARRAD AM, PAPARELLA AS, et al. Heterocyclic acyl-phosphate bioisostere-based inhibitors of Staphylococcus aureus biotin protein ligase[J]. Bioorg Med Chem Lett, 2014, 24 (19): 4689.
[22] POLYAK SW, BAILEY LM, AZHAR A, et al. Biotin (Vitamin H or B7)in: Micronutrients: Sources, properties and health benefits[M]. Nova Science Publishers, New York, 2012: 65.
[23] ZEMPLENI J, WIJERATNE SS, HASSAN YI. Biotin[J]. Biofactors Oxf Engl, 2009, 35(1): 36.
[24] ROJE S. Vitamin B biosynthesis in plants[J]. Phytochemistry, 2007, 68(14): 1904.
[25] AZHAR A, POLYAK SW, BOOKER GW. Mechanisms of biotin transport[J]. Biochem Anal Biochem, 2015, 04(04): 210.
[26] HEBBELN P, RODIONOV DA, ALFANDEGA A, et al. Biotin uptake in prokaryotes by solute transporters with an optional ATP-binding cassette-containing module[J]. Proc Natl Acad Sci U S A, 2007, 104(8): 2909.
[27] RODIONOV DA, MIRONOV AA, GELFAND MS. Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea[J]. Genome Res, 2002, 12(10): 1507.
[28] SALAEMAE W, YAP MY, WEGENER KL, et al. Nucleotide triphosphate promiscuity in Mycobacterium tuberculosis dethiobiotin synthetase[J]. Tuberculosis (Edinb), 2015, 95(3): 259.
[29] WOONG PARK S, KLOTZSCHE M, WILSON DJ, et al. Evaluating the sensitivity of Mycobacterium tuberculosis to biotin deprivation using regulated gene expression[J]. PLoS Pathog, 2011, 7(9): e1002264.
[30] SASSETTI CM, RUBIN EJ. Genetic requirements for mycobacterial survival during infection[J]. Proc Natl Acad Sci U S A, 2003, 100(22): 12989.
[31] KEER J, SMEULDERS MJ, GRAY KM, et al. Mutants of *mycobacterium smegmatis* impaired in stationary-phase survival[J]. Microbiology, 2000, 146(Pt 9): 2209.
[32] SASSETTI CM, BOYD DH, RUBIN EJ. Comprehensive identification of conditionally essential genes in mycobacteria[J]. Proc Natl Acad Sci U S A, 2001, 98(22): 12712.
[33] RENGARAJAN J, BLOOM BR, RUBIN EJ. Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages[J]. Proc Natl Acad Sci U S A, 2005, 102 (23): 8327.
[34] YU J, NIU C, WANG D, et al. MMAR_2770, a new enzyme involved in biotin biosynthesis, is essential for the growth of Mycobacterium marinum in macrophages and zebrafish[J]. Microbes Infect, 2011, 13(1): 33.
[35] TOMMASI R, BROWN DG, WALKUP GK, et al. ESKAPEing the labyrinth of antibacterial discovery[J]. Nat Rev Drug Discov, 2015, 14(8): 529.
[36] COOPER MA. A community-based approach to new antibiotic discovery[J]. Nat Rev Drug Discov, 2015, 14(9): 587.
[37] LING LL, SCHNEIDER T, PEOPLES AJ, et al. A new antibiotic kills pathogens without detectable resistance[J]. Nature, 2015, 517(7535): 455.

**Received**: 2016－03－02　**Editor**: Peng Hai-yan



**作者合影**：第一排从左到右为李燕，Sternicki Louise M，Wegener Kate L；第二排从左到右为 Polyak Steven W，Booker Grant W，Thompson Andrew P。于 2016 年 4 月 20 日摄于澳大利亚阿德莱德大学。

Booker Grant W is professor of biochemistry at the University of Adelaide, Australia. Together with Dr Polyak Steven W and Dr Wegener Kate L, he runs an antibacterial drug discovery program with a focus on new treatments for Staphylococcus aureus and Mycobacterium tuberculosis. Their approach is exploit enzymes that require the important micronutrient biotin as drug targets for the discovery of novel classes of antibiotics. The team specializes in structure guided drug design and fragment based drug discovery. Thompson Andrew P is a PhD candidate working on biotin biosynthetic enzymes as a new drug target for TB. Sternicki Louise M's graduate project is investigating biotin protein ligase as a drug target for new anti-infectives. Dr Li Yan was the recipient of a Jiangsu Health International Exchange Program sponsorship to work on an early stage drug discovery at the University of Adelaide.

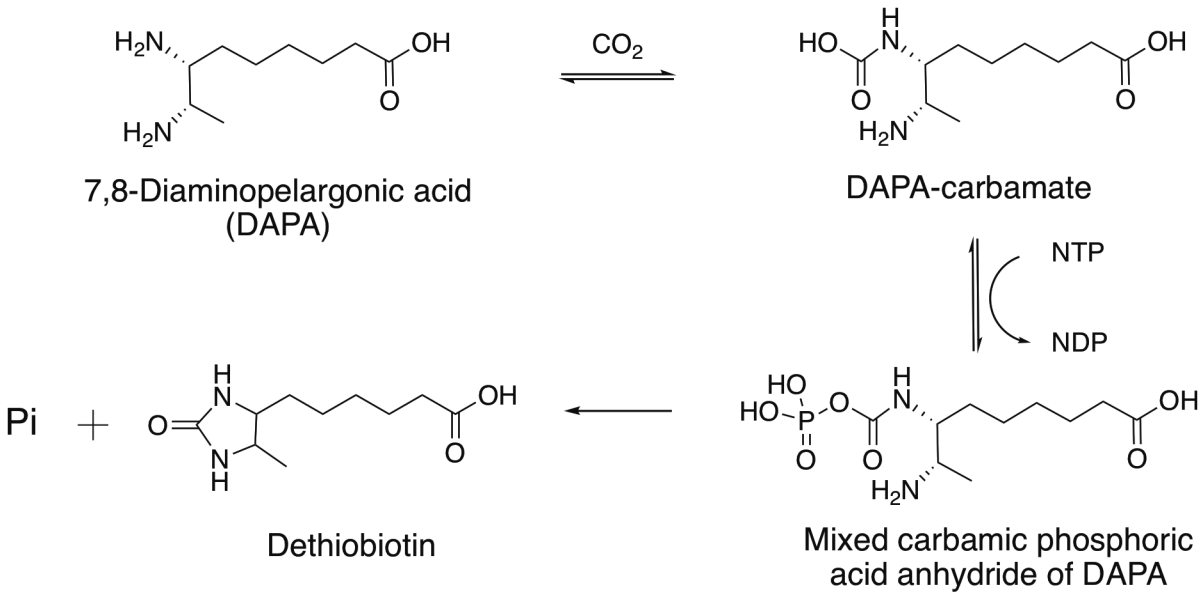
Booker Grant W 教授，澳大利亚阿德莱德大学生物化学系教授，与 Polyak Steven W 博士和 Wegener Kate L 博士组成科研团队，共同承担治疗金黄色葡萄球菌和结核分枝杆菌的新型抗生素研发项目。研究主要以生物素相关酶类作为药物靶点筛选新型抗生素，生物素是微生物的重要生长因子，正常代谢不可或缺的关键物质。该研究团队擅长药物化学结构改造和小分子新药研发工作。Thompson Andrew P 为在读博士研究生，主要从事以生物素合成酶为靶点的新型药物筛选工作。Sternicki Louise M 为在读博士研究生，主要从事以生物素蛋白连接酶为靶点的新型药物筛选工作。李燕博士为江苏卫生国际交流项目资助的访问学者，主要从事结核病诊断方法和结核病防治相关研究；现在阿德莱德大学从事以生物素蛋白连接酶为靶点的新型抗生素早期研究的相关工作。

## 1.4 Dethiobiotin Synthetase

As demonstrated in the published review, there is considerable evidence to support targeting biotin biosynthesis in anti-tubercular drug design efforts. Of particular interest is the demonstration that the biotin biosynthesis is essential for the acute and dormant phases of *M. tuberculosis* infection (Woong Park et al., 2011). The work presented in this thesis concerns *Mycobacterium tuberculosis* dethiobiotin synthetase (*Mt*DTBS), the enzyme responsible for catalysing the penultimate step of biotin biosynthesis (Dey *et al*., 2010; Salaemae *et al*., 2015). This enzyme was determined to be preferable to other enzymes in the pathway for drug discovery purposes. Firstly, there are two genes that produce differing isoforms of the initial enzyme in the pathway: 7-keto-8-aminopelargonic acid synthase (KAPAS), making this a challenging target for drug discovery as both forms would need to be simultaneously inhibited to arrest biotin biosynthesis activity (Salaemae *et al*., 2016). Secondly, the final enzyme in the pathway: biotin synthase, exhibits very low catalytic rates and is more functionally similar to a substrate than an enzyme (Salaemae *et al*., 2016). This suggests that high doses of inhibitor would be required to completely inhibit this step. Finally, it was previously found that 7,8-diaminopelargonic acid synthase (DAPAS) did not overexpress well using the facilities at the University of Adelaide. Consequently, *Mt*DTBS was the deemed the best candidate for an inhibitor development project.

*Mt*DTBS catalyses energy dependent conversion of 7,8-diaminopelargonic acid (DAPA) to dethiobiotin (DTB; Scheme 1) (Huang *et al*., 1995; Käck *et al*., 1998). Unusually, and in contrast to homologues from other species, *Mt*DTBS can utilize a broad range of nucleoside triphosphates (NTPs) to carry out its reaction, namely; adenosine, cytidine, guanosine, inosine, thymidine and uridine triphosphate (Salaemae *et al*., 2015). It was determined that *Mt*DTBS exhibits a slight $k_{cat}$ preference for cytidine triphosphate (CTP $k_{cat} = 15.2 \pm 0.3 \times 10^{-3}$ $s^{-1}$, other NTPs $k_{cat} = 10.2 - 11.8 \times 10^{-3}$ $s^{-1}$), despite similar $K_m$ values (CTP $K_m = 25.2 \pm 1.7$ μM, other NTPs $K_m = 17 - 30$ μM) (Salaemae *et al*., 2015). While a crystal structure with CTP revealed how *Mt*DTBS bound this substrate (PDB ID: 4WOP), the mechanism behind NTP promiscuity had eluded structural characterisation. The manuscript presented in Chapter 3 contains insights into this phenomenon via X-ray crystallographic investigation.

*Mt*DTBS has previously been investigated by X-ray crystallography (Dey *et al*., 2010; Salaemae *et al*., 2015). The overall fold consists of alternating α-helices and β-sheets separated by interconnecting coil regions. The enzyme is only biologically active as a homo-dimer and presents as a dimer of dimers in the crystallographic asymmetric unit (Salaemae et al., 2015). Crystal structures of *Mt*DTBS in complex with several different ligands have been solved (Table 1), revealing that the active site contains adjacent CTP and DAPA binding pockets (Figure 1.1). While the overall fold is similar, the cytidine binding pocket differs from that in DTBS from other species, of which crystal structures have been solved in complex with multiple ligands (Table 1). The impact of these variations is discussed in the manuscript in Chapter 3.




**Scheme 1:** The reaction catalysed by *Mt*DTBS. After binding to *Mt*DTBS, 7,8-diaminopelargonic acid (DAPA) is converted to DAPA-carbamate by reaction with a carbon dioxide molecule. A nucleoside triphosphate (NTP) catalyses phosphoryl transfer to DAPA, resulting in a mixed carbamic phosphoric acid DAPA anhydride (Käck *et al.*, 1998). This molecule then collapses to close the ureido ring of dethiobiotin, releasing inorganic phosphate (Pi) in the process. Figure adapted from the manuscript in Chapter 3.

Despite the previous structural characterisation of *Mt*DTBS, there are no published examples of inhibitors developed against this enzyme, unlike its *E. coli* counterpart (Rendina *et al.*, 1999). Given the wealth of information demonstrating that biotin biosynthesis is an anti-tubercular drug target, this project aimed to identify and develop small molecules that would bind with high affinity to *Mt*DTBS. During development of the crystallographic technique (Chapter 2) and characterization of the *Mt*DTBS NTP utilization mechanism (Chapter 3), several positions within the active site were identified as crucial binding regions for substrates. This information provided useful guidance for lead molecule development (Chapter 4).

**Table 1:** X-ray crystallographic structures of DTBS enzymes from the protein data bank (PDB).

| **Organism** | **Ligand bound** | **PDB ID** |
|---|---|---|
| *Mycobacterium tuberculosis* | Apo (unliganded) | 3FGN (Dey *et al*., 2010) |
| | Dethiobiotin, $Mg^{2+}$, $PO_4^{2-}$ | 3FPA (Dey *et al*., 2010) |
| | DAPA-carbamate | 3FMF (Dey *et al*., 2010) |
| | KAPA, $SO_4^{2-}$ | 3FMI (Dey *et al*., 2010) |
| | CTP, $SO_4^{2-}$ | 4WOP (Salaemae *et al*., 2015) |
| *Escherichia coli* | Apo | 1BYI (Sandalova *et al*., 1999) |
| | Apo | 1DTS (Huang *et al*., 1995) |
| | ATP, DAPA | 1A82 (Käck *et al*., 1998a) |
| | ADP | 1DAD (Huang *et al*., 1995) |
| | ADP, Mixed carbamic phosphoric acid anhydride of DAPA | 1DAK (Käck *et al*., 1998a) |
| | ADP, 8-amino-7-carboxyamino-nonanoic acid (DAPA) with aluminium fluoride | 1BS1 (Käck *et al*., 1998b) |
| | ADP, dethiobiotin, phosphate | 1DAM (Käck *et al*., 1998b) |
| | ADP, 7-(carboxyamino)-8-amino-nonanoic acid (DAPA-carbamate) | 1DAF (Huang *et al*., 1995) |
| | 3-(1-aminoethyl)-nonanedioic acid | 1DAE (Huang *et al*., 1995) |
| | DAPA-carbamate, 5'-adenosyl-methylene-triphosphate | 1DAG (Huang *et al*., 1995) |
| | DAPA, 5'-adenosyl-methylene-triphosphate | 1DAH (Huang *et al*., 1995) |
| | DAPA-carbamate | 1DAI (Huang *et al*., 1995) |
| | Sulfate | 1DBS (Alexeev *et al*., 1994) |
| *Helicobacter pylori* | Apo | 2QMO (Porebski *et al*., 2012) |
| | ATP | 3QXC (Porebski *et al*., 2012) |
| | ATP, 8-aminocaprylic acid, phosphate | 3MLE (Porebski *et al*., 2012) |
| | ADP, 8-aminocaprylic acid | 3QXH (Porebski *et al*., 2012) |
| | ANP | 3QXS (Porebski *et al*., 2012) |
| | GTP | 3QXJ (Porebski *et al*., 2012) |
| | GDP | 3QY0 (Porebski *et al*., 2012) |
| | GDP, 8-aminocaprylic acid | 3QXX (Porebski *et al*., 2012) |
| *Franciscella tularensis* | Apo | 3OF5 (To be published) |

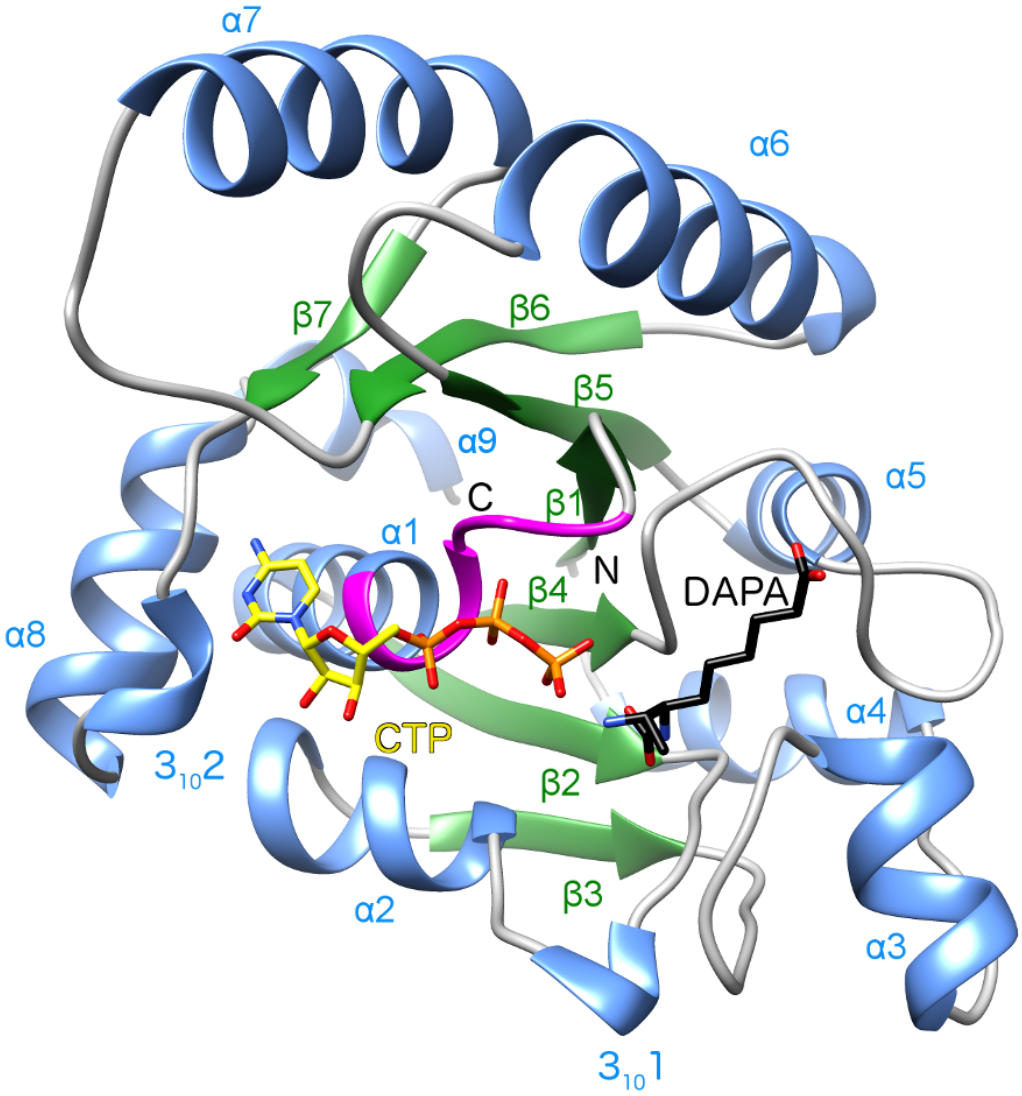



**Figure 1.1:** An overlay of CTP and DAPA-carbamate bound *Mt*DTBS crystal structures (PDB IDs: 4WOP, 3FMF, respectively). α-helices and $3_{10}$ helices are coloured in blue, β-sheets in green, interconnecting coil in grey and the phosphate binding loop in magenta. A monomer is shown for clarity, the homo-dimerisation interface is formed by α3, α5 and α6. The cytosine moiety of CTP binds to a pocket formed by the coils connecting β6-α7 and β7-$3_{10}2$. The triphosphate moiety of CTP binds to the phosphate binding loop. DAPA-carbamate binds adjacent to CTP in a deep pocket that is formed at the dimerization interface. Figure adapted from the manuscript in Chapter 3.

## 1.5 Approaches to antibiotic drug discovery

Resistance to current antibiotics poses a severe threat to global health. One method to combat resistance is the development of new antibiotics, a practice that has experienced a downturn in recent decades. This is the result of decreased investments, but also a failure in current drug discovery methods. A review from AstraZeneca extensively detailed their antibiotic high-throughput screening (HTS) programs from 2001–2010, from which no drugs were produced (Tommasi et al., 2015). HTS utilizes extremely high throughput assays to detect inhibition of a purified protein target using very large libraries (~$10^6$) of complex, drug like compounds. This review highlighted many issues with current antibiotic drug discovery strategies. Of particular note were the high rates of attrition when taking molecules with high protein target potency to bacterial cell-based assays. This was due to various medicinal chemistry issues, including poor chemical tractability of hits, resulting from the large, complex and drug-like molecules that comprised the compound library. It was also hypothesized that there was a lack of understanding of the biological systems involved, such as bacterial membrane permeability. Indeed, antibiotic drug discovery faces additional challenges when compared to drugs targeting human diseases, as both bacterial and human metabolism must be considered, as well as differing cell wall and membrane structures. *M. tuberculosis*, for example, possesses a unique cell wall structure made up of mycolic acids and persists within alveolar macrophages during chronic infection. Anti-mycobacterial therapeutics must therefore have the appropriate properties to not only gain entry into both cell types, but also retain function during macrophage and mycobacterial metabolism. New innovations are required to overcome the difficulties currently faced by antibiotic and anti-mycobacterial drug discovery.

Fragment-based lead discovery (FBLD) is considered a promising alternative to HTS (Erlanson *et al*., 2016; Mashalidis *et al*., 2013). In a FBLD campaign, a small library (~$10^3$) of relatively simple molecules is screened against protein targets to generate a number of low affinity hits ($K_D$ > 500 μM) (Baker, 2013; Joseph-McCarthy *et al*., 2014). Smaller compound libraries are easier to handle, making FBLD campaigns more attractive and accessible to a wide range of researchers, particularly academic groups (Erlanson *et al*., 2016). The small size of fragments results in more efficient sampling of chemical space than HTS, increasing the likelihood of success (Erlanson *et al*., 2016; Lamoree and Hubbard, 2018). Unlike HTS hits, which can be difficult to modify, fragment hits are highly tractable, and are routinely developed into larger and more complex molecules that are more potent inhibitors of the protein target. This process involves cyclical rounds of development and testing with structural, biophysical and biochemical methods. As this is an iterative process, this can coincide with observation and optimisation of the pharmacological properties of the compounds (Erlanson *et al*., 2016). It is hypothesized that this would be beneficial for antibiotic drug discovery, where properties such as bacterial metabolism and membrane permeability could be optimised during compound design. The work presented in Chapter 4 entails the continuation of a FBLD campaign against *Mt*DTBS; the validation of fragment binding and preliminary structurally guided compound elaboration.

A central theme to this thesis is that structural, biophysical and biochemical characterisation of the target protein will greatly aid the process of fragment development. This is demonstrated again in the review by AstraZeneca (Tommasi *et al*., 2015). Here it was revealed that HTS against biotin ligase (BirA) uncovered no compounds with reproducible inhibitory activity in orthogonal assays. BirA has been shown to utilize an ordered catalytic mechanism initiated by binding biotin, which induces closure of the ‘biotin binding loop’ to

create the site to which adenosine triphosphate binds (Sternicki *et al*., 2017). It is likely that HTS was not able to identify compounds that inhibited BirA because this mechanism was not considered; these compounds did not outcompete biotin for binding to the highly specific biotin binding pocket, and in the absence of biotin, the adenosine triphosphate binding site would have been inaccessible. However, with consideration of the ordered catalytic mechanism of BirA, the Booker and Abell Laboratories at the University of Adelaide utilized a structural and biochemically guided approach that resulted in the generation of several lead compound series against the protein (Feng *et al*., 2016; Paparella *et al*., 2018a, 2018b; Soares da Costa *et al*., 2012; Tieu *et al*., 2015). These lead compounds mimicked the reaction intermediate, possessing a biotin moiety conjugated to alternative adenosine analogues. These compounds, designed with consideration of the structural and mechanistic properties of catalysis, inhibited the enzyme with high efficacy. This highlights that structural and mechanistic knowledge of enzyme catalysis can be crucial for ligand identification and design.

As previously mentioned, the accessibility of FBLD can be taken advantage of by smaller and more focussed groups, such as academic labs that have expertise with particular enzymes. The Booker laboratory has researched biotin biology extensively, publishing over 30 articles on the subject since 2008. In Chapters 2 and 3, these studies have been extended to investigations into the detailed ligand-binding characteristics of *Mt*DTBS, revealing structural and biophysical details of interactions at the cytidine binding site, phosphate binding loop, and DAPA pocket. The resulting information guided the preliminary compound identification and design presented in Chapter 4 and will continue to guide future inhibitor development. Altogether, the work presented in this thesis extends our knowledge of *Mt*DTBS structure and function and represents a crucial progression towards the development of inhibitors of *Mt*DTBS, and ultimately *Mycobacterium tuberculosis*, in order to provide new avenues of treatment for antibiotic-resistant tuberculosis.

## 1.5 Project Aims

This study aimed to continue the investigation of *Mt*DTBS, an enzyme that has been identified as a promising target for the development of anti-tuberculosis therapies.

The first aim was to develop and optimize protocols for the structural and biophysical of *Mt*DTBS. With protocols for the expression and purification of *Mt*DTBS, as well as a biochemical activity assay available to the Booker Lab, additional protocols would allow for comprehensive study of this enzyme.

The second aim was to utilize the developed protocols to structurally and biophysically characterize *Mt*DTBS promiscuous NTP utilization during catalysis. As this is an unusual phenomenon, uncovering further insight into this capability would be of interest to the scientific community.

The third aim of this study was to assist in the development of compounds which would bind to *Mt*DTBS with high affinity, with the ultimate aim of inhibiting the activity of this enzyme. This aim was influenced by the results of the first two aims, requiring the development of optimized protocols, as well as insight into the structural and biophysical nature of *Mt*DTBS ligand binding.

# Chapter 2:

Publication: Precipitant-ligand exchange technique reveals ADP binding mode in *Mt*DTBS

## 2.1 Statement of Authorship

| | |
|---|---|
| Title of Paper | Precipitant-ligand exchange technique reveals ADP binding mode in *Mt*DTBS |
| Publication Status | Published |
| Publication Details | Thompson, A. P., Wegener, K. L., Booker, G. W., Polyak, S. W. & John B. Bruning (2018). *Acta Cryst. D***74**. https://doi.org/10.1107/S2059798318010136. |

## Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Andrew P. Thompson | | |
| Contribution to the Paper | Performed all experimental work and analysis, including expression of *Mt*DTBS enzyme, crystallographic experiments including structure solving and interpretation. Prepared figures and manuscript. | | |
| Overall percentage (%) | 80% | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature | | Date | 20/11/18 |

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| | | | |
|---|---|---|---|
| Name of Co-Author | Kate L. Wegener | | |
| Contribution to the Paper | Assisted in experimental design, manuscript preparation and proof-reading. | | |
| Signature | | Date | 5/12/18 |

| Name of Co-Author | Grant W. Booker | | |
|---|---|---|---|
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 20/11/2018 |

| Name of Co-Author | Steven W. Polyak | | |
|---|---|---|---|
| Contribution to the Paper | Conceived *Mt*DTBS project. Assisted in experimental design, manuscript preparation and proof-reading. | | |
| Signature | | Date | 23/11/2018 |

| Name of Co-Author | John B. Bruning | | |
|---|---|---|---|
| Contribution to the Paper | Conceived and guided crystallographic experiments, helped with data processing and interpretation, assisted manuscript preparation and proof-reading. Corresponding author. | | |
| Signature | | Date | 20/11/2018 |

STRUCTURAL BIOLOGY
ISSN 2059-7983

# Precipitant–ligand exchange technique reveals the ADP binding mode in *Mycobacterium tuberculosis* dethiobiotin synthetase

**Andrew P. Thompson,[a] Kate L. Wegener,[b] Grant W. Booker,[a] Steven W. Polyak[a] and John B. Bruning[b]***

[a]Molecular and Biomedical Science, The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia, and [b]Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia. *Correspondence e-mail: john.bruning@adelaide.edu.au

Received 22 May 2018
Accepted 13 July 2018

Edited by B. Kobe, University of Queensland, Australia

**Keywords:** dethiobiotin synthetase; *Mycobacterium tuberculosis*; crystallography; precipitant; nucleoside polyphosphate; ligand; exchange; soak.

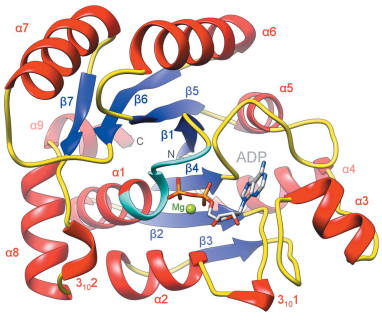
**PDB references**: ADP–*Mt*DTBS, 6czd; CTP–*Mt*DTBS, 6e05; 6e06

**Supporting information**: this article has supporting information at journals.iucr.org/d

Dethiobiotin synthetase from *Mycobacterium tuberculosis* (*Mt*DTBS) is a promising antituberculosis drug target. Small-molecule inhibitors that target *Mt*DTBS provide a route towards new therapeutics for the treatment of antibiotic-resistant tuberculosis. Adenosine diphosphate (ADP) is an inhibitor of *Mt*DTBS; however, structural studies into its mechanism of inhibition have been unsuccessful owing to competitive binding to the enzyme by crystallographic precipitants such as citrate and sulfate. Here, a crystallographic technique termed precipitant–ligand exchange has been developed to exchange protein-bound precipitants with ligands of interest. Proof of concept for the exchange method was demonstrated using cytidine triphosphate (CTP), which adopted the same binding mechanism as that obtained with traditional crystal-soaking techniques. Precipitant–ligand exchange also yielded the previously intractable structure of *Mt*DTBS in complex with ADP solved to 2.4 Å resolution. This result demonstrates the utility of precipitant–ligand exchange, which may be widely applicable to protein crystallography.

## 1. Introduction

Biotin biosynthesis has been identified as a promising pathway for the development of new antituberculosis chemotherapies required to combat the rise of multidrug-resistant infections (Salaemae *et al.*, 2011, 2016; Thompson *et al.*, 2016; Dai *et al.*, 2014, 2015; Dey *et al.*, 2010; Woong Park *et al.*, 2011). The penultimate reaction in the pathway is the energy-dependent conversion of 7,8-diaminopelargonic acid (DAPA) to dethiobiotin (DTB), catalysed by dethiobiotin synthetase (DTBS). Small molecules that target DTBS represent valuable starting points for antitubercular drug discovery. We have previously demonstrated that ADP is an inhibitor of *Mycobacterium tuberculosis* DTBS (*Mt*DTBS; $K_i$ = 0.14 m*M* with respect to ATP) and binds with comparable affinity to ATP (ADP $IC_{50}$ = 248 µ*M*, ATP $IC_{50}$ = 349 µ*M*) in a competitive displacement assay (Salaemae, 2015). Thus, ADP provides a potential starting point for the chemical development of nucleoside analogues to treat antibiotic-resistant tuberculosis. The crystal structure of *Mt*DTBS in complex with cytidine triphosphate (CTP; PDB entry 4wop) revealed the binding mode for this ligand (Salaemae *et al.*, 2015). The cytosine moiety bound in a pocket formed by the coils connecting $\beta$7 to $3_{10}$2 and $\beta$6 to $\alpha$7 (Figs. 1 and 2), and the polyphosphate group bound at the phosphate-binding loop (P-loop, Walker A motif; depicted in Fig. 1). However, efforts to crystallize *Mt*DTBS in complex with other nucleosides such as ADP have been unsuccessful



(Salaemae *et al.*, 2015; Dey *et al.*, 2010) despite similar structures having been obtained with homologues from other bacteria (Huang *et al.*, 1995; Käck *et al.*, 1998). In our repeated attempts to solve the ADP–*Mt*DTBS crystal structure, electron density consistent with a sulfate ion bound at the P-loop was observed, resulting from the high concentrations of ammonium sulfate in the crystallization condition (1.2–1.7 *M* ammonium sulfate, 0.1 *M* Tris pH 8, 10–15% glycerol; adapted from Salaemae *et al.*, 2015). Concordant with this, a sulfate ion was observed instead of CTP in one active site of the CTP–*Mt*DTBS crystal structure (Salaemae *et al.*, 2015). We propose that precipitants that are present at high concentrations in the crystallization condition directly compete with ADP for binding to the P-loop. Consequently, we sought to overcome precipitant occupation of the active site in order to investigate *Mt*DTBS in complex with ADP.

Crystallographic reagents commonly bind to distinct regions of protein molecules. Table 1 depicts the prevalence of common crystallographic reagents modelled in structures deposited as Protein Data Bank (PDB) entries. Of the total submissions in the PDB (139 357 at the time of writing), sulfate is present in 12.4% of submissions and glycerol (a common cryoprotectant) is present in 10.5%. These two molecules are the most commonly modelled free ligands in the PDB, indicating that this is a common occurrence in protein crystallography. Efforts to avoid crystallographic reagents may require either a new crystallization condition or the use of a homologous protein. Alternatively, techniques to remove these compounds by crystal manipulation and soaking have also been reported. Examples of such approaches in the literature include crystal washing to remove dioxane, back-soaking to remove substrates and soaking with homologous

**Table 1**
The prevalence of selected crystallographic reagents modelled as free ligands in Protein Data Bank entries at the time of writing (139 357 total structures).

| Crystallographic molecule | No. of entries | Percentage of total submissions |
|---|---|---|
| Sulfate ion | 17338 | 12.4 |
| Glycerol | 14594 | 10.5 |
| Ethane-1,2-diol (EDO) | 8370 | 6.0 |
| Phosphate ion | 5046 | 3.6 |
| Acetate ion | 4334 | 3.1 |
| Di(hydroxyethyl)ether (nonpolymer PEG) | 2444 | 1.8 |
| Citrate | 409 | 0.3 |

precipitants (for example, replacing citrate with malonate) (reviewed in Öster *et al.*, 2015). Ligand-exchange methods have previously been utilized to substitute co-crystallized peptides (Kontopidis *et al.*, 2003) and degraded substrates (Ciccone *et al.*, 2015) with ligands. In the current study, this principle was applied to exchange active-site-bound crystallographic precipitants with ligands. This alternate methodology was required to ensure the stability of *Mt*DTBS crystals, which was reliant upon a high concentration of a suitable compound (either precipitant or ligand) being bound within the P-loop.

Here, we describe a soaking methodology to overcome the occlusion of ligands of interest by crystallographic precipitant molecules in the active site. Having ascertained that alternate crystal-growth conditions also resulted in precipitants bound to the P-loop, we assessed the crystal integrity in varying concentrations of the precipitants sulfate and citrate, and ultimately replaced the precipitant with high concentrations of the ligand of interest. As proof of concept, two structures of *Mt*DTBS in complex with CTP were solved using precipitant–ligand exchange with crystals from two independent growth conditions. In both structures, CTP adopted the same binding pose as demonstrated by our previous crystal structure (PDB entry 4wop). Subsequently, a crystal structure of the previously intractable complex with the inhibitor ADP was obtained. We suggest that our precipitant–ligand exchange technique may be broadly applicable to crystallography of protein–ligand complexes and may be especially important for obtaining crystals of weak-affinity complexes such as those studied in fragment-based drug discovery.

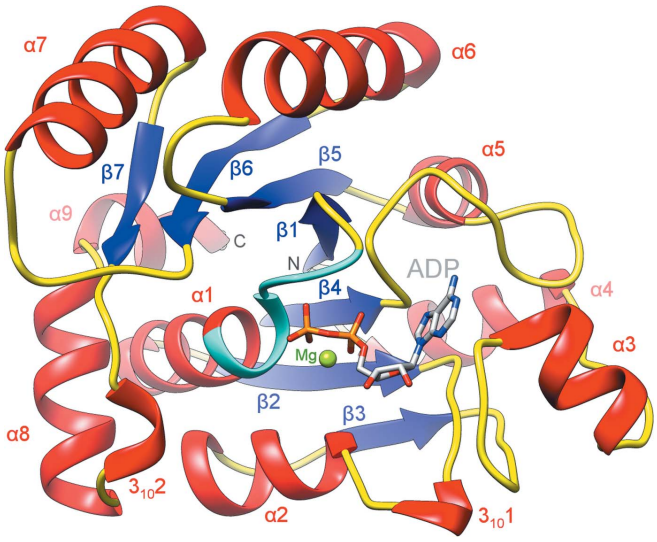



**Figure 1**
The ADP-bound monomer of *Mt*DTBS (chain *A*; PDB entry 6czd). Secondary structure is depicted as ribbons: helices are coloured red, $\beta$-sheets blue and interconnecting coils yellow. ADP (grey sticks) bound distal to the cytidine-binding site (coils connecting $\beta7$ to $3_{10}2$ and $\beta6$ to $\alpha7$). The $\beta$-phosphate bound at the P-loop (cyan) coordinated to a magnesium ion (green sphere).

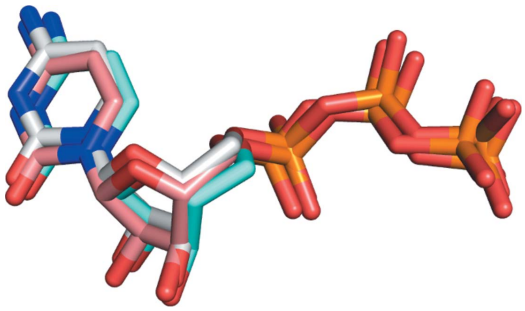


**Figure 2**
Superposition of CTP in our previously published structure (white sticks; PDB entry 4wop) and those solved by precipitant–ligand exchange (pink, sulfate-grown crystal; blue, citrate-grown crystal).

## 2. Materials and methods

### 2.1. Protein expression and purification

*Mt*DTBS was expressed and purified as described previously (Salaemae *et al.*, 2015), except that a HiPrep 26/60 Sephacryl S-300 HR column (GE Healthcare) was used as an additional purification step following IMAC chromatography. *Mt*DTBS was concentrated to 10 mg ml$^{-1}$ and stored in the following solution: 25 m*M* Tris pH 7.5, 30 m*M* NaCl, 1 m*M* EDTA, 1 m*M* dithiothreitol (DTT), 5%(*v*/*v*) glycerol.

### 2.2. Nucleoside polyphosphate preparation

Nucleoside polyphosphates (NpPs) were purchased as analytical grade powders (Sigma–Aldrich) and solubilized to 100–500 m*M* in Milli-Q $H_2O$ followed by adjustment to pH 7 with NaOH.

### 2.3. General crystallographic methods

All *Mt*DTBS crystals were grown *via* the hanging-drop method with a well volume of 500 µl and a drop ratio of 1:1 protein:well solution at 16°C unless otherwise stated. Sulfate ion-bound *Mt*DTBS crystals were grown in an ammonium sulfate crystallization condition consisting of 1.2–1.7 *M* ammonium sulfate, 0.1 *M* Tris pH 8, 10–15% glycerol. Citrate-bound *Mt*DTBS crystals were grown in a citrate crystallization solution consisting of 0.7 *M* sodium citrate, 0.1 *M* imidazole pH 7.0 with 20% glycerol.

### 2.4. Serial dilutions and ligand-exchange solutions

Serial Dilution Solution 1: a twofold serial dilution of citrate (0.35–0.011 *M*) was produced by mixing the citrate crystallization condition with an equal volume of precipitant-free solution (0.1 *M* imidazole pH 7.0 with 20% glycerol).

Serial Dilution Solution 2: a twofold serial dilution of citrate (0.35–0.011 *M*) in the presence of 5 m*M* ATP was produced by mixing the citrate crystallization condition with a solution containing ATP in place of citrate (10 m*M* ATP, 1 m*M* $MgCl_2$, 0.1 *M* imidazole pH 7.0, 20% glycerol).

Exchange Solution 1 consisted of 300 m*M* ATP, 60 m*M* $MgCl_2$, 0.022 *M* sodium citrate, 0.1 *M* imidazole pH 7.0, 20% glycerol.

Exchange Solution 2 consisted of 300 m*M* ATP, 30 m*M* $MgCl_2$, 0.1 *M* Tris pH 8, 10–20% glycerol. The well solution consisted of 0.1 *M* Tris pH 8, 10–20% glycerol.

Exchange Solution 3 consisted of 100 m*M* CTP, 10 m*M* $MgCl_2$, 0.1 *M* imidazole pH 7, 20% glycerol.

Exchange Solution 4 consisted of 100 m*M* CTP, 10 m*M* $MgCl_2$, 0.1 *M* Tris pH 8, 25% glycerol.

Exchange Solution 5 consisted of 70 m*M* ADP, 30 m*M* $MgCl_2$, 0.1 *M* Tris pH 8, 10–20% glycerol. The well solution consisted of 0.1 *M* Tris pH 8, 10–20% glycerol.

### 2.5. Data collection and processing

All crystals were flash-cooled in liquid nitrogen and subjected to X-ray diffraction on the MX1 beamline at the Australian Synchrotron, which is part of ANSTO (McPhillips *et al.*, 2002). 360 images were collected with 1° oscillation at 100 K and at a wavelength of 0.9537 Å. Data were indexed, scaled and merged using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (*CCP*4; Winn *et al.*, 2011). Resolution truncation was performed in accordance with previously described $CC_{1/2}$ cutoff values (Karplus & Diederichs, 2012, 2015). Phasing was performed with *Phaser-MR* (McCoy *et al.*, 2007) using an in-house search model. The structure was iteratively refined with manual modelling in *Coot* (Emsley & Cowtan, 2004) and refinement with *phenix.refine* (Adams *et al.*, 2010) until the *R* factors converged and significant difference density was accounted for. The structure was validated using *MolProbity* as part of the *PHENIX* package. Structure figures were generated with the *UCSF Chimera* package (Pettersen *et al.*, 2004) and *MacPyMOL* (Schrödinger). Coordinates and structure factors were deposited in the Protein Data Bank for ADP–*Mt*DTBS (PDB entry 6czd), CTP–*Mt*DTBS from crystals grown in ammonium sulfate (PDB entry 6e05) and in sodium citrate (PDB entry 6e06).

## 3. Results

### 3.1. Alternate crystal conditions consistently contained precipitants bound at the P-loop

In order to obtain crystals of nucleoside polyphosphates (NpPs) bound to *Mt*DTBS without a competing precipitant, alternate crystallization conditions were explored. To avoid the use of ammonium sulfate, *Mt*DTBS crystals were grown in a citrate crystallization solution adapted from Dey *et al.* (2010) (see §2). However, no NpP-complexed crystal structures were obtained *via* either co-crystallization or soaking. Instead, electron density resembling citrate was observed bound to the P-loop in all solved crystal structures (data not shown). Further crystallographic screening yielded hits in conditions containing high concentrations of sulfate, citrate or organic acid mixtures (Tacsimate), all of which are likely to bind at the same location. Therefore, a technique was developed to replace the precipitants with an NpP.

### 3.2. Reducing the precipitant concentration caused crystal instability

The stability of *Mt*DTBS crystals in decreasing concentrations of precipitant was assessed. *Mt*DTBS crystals were grown in the citrate crystallization condition before being transferred directly into each solution of a twofold serial dilution of citrate (0.35–0.011 *M* sodium citrate; Serial Dilution Solution 1; §2). As expected, crystal stability correlated with the precipitant concentration (Fig. 3). Crystals transferred into solutions containing less than 0.044 *M* citrate dissolved rapidly (<30 min), those in 0.088 *M* citrate dissolved overnight and those in 0.175 *M* citrate cracked severely. It was therefore hypothesized that crystal integrity was dependent on the presence of a suitable molecule, in this case the precipitant, bound to the P-loop of *Mt*DTBS.

### 3.3. The addition of ligand increased the stability of crystals

As NpP molecules also bind to the P-loop, it was proposed that these molecules would rescue crystal stability in the absence of precipitant. To test this, crystal stability was assessed in a twofold serial dilution series of citrate containing a constant concentration of 5 m*M* ATP (Serial Dilution Solution 2; §2). There was no appreciable improvement in crystal stability using this concentration of ligand (Table 2). It was hypothesized that the percentage of *Mt*DTBS molecules in the crystal bound to a ligand (precipitant or NpP) was not high enough to retain crystal integrity. Therefore, much higher concentrations of NpP were trialled, analogous to the high concentrations of precipitant required for crystal growth described previously. Crystals were transferred into a solution of 0.022 *M* citrate that was supplemented with 300 m*M* ATP (Exchange Solution 1; §2). The crystals of *Mt*DTBS showed improved stability, as they did not degrade overnight. Similarly, crystals grown in sulfate were stable when supplemented with ATP. This was consistent with the hypothesis that crystal integrity was dependent on the presence of a suitable ligand to occupy the P-loop.

To ensure that the ligand of interest would be unobscured by precipitant density in the resulting crystal structure, efforts were made to completely remove the protein-bound precipitant. *Mt*DTBS crystals were grown in the ammonium sulfate crystallization condition before being transferred into an 8 µl drop of Exchange Solution 2 containing 300 m*M* ATP. An 8 µl drop size was used to minimize dilution by any of the original buffer components in the transfer process. This resulted in crystals

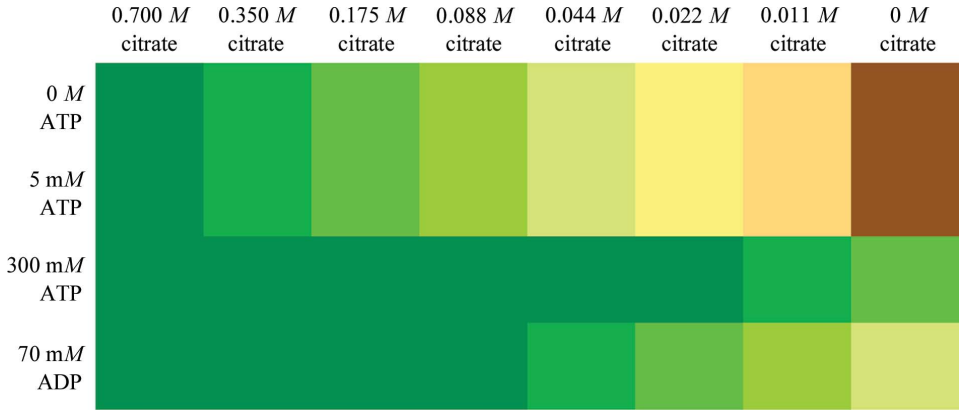



**Figure 3**
Qualitative crystal stability after transfer into solutions containing varying concentrations of citrate precipitant and ATP or ADP. Crystal stability is denoted from highest to lowest from green to brown.

**Table 2**
Data-collection and processing statistics for *Mt*DTBS crystal structures.

Values in parentheses are for the outer shell.

| | CTP–*Mt*DTBS (sulfate), PDB entry 6e05 | CTP–*Mt*DTBS (citrate), PDB entry 6e06 | ADP–*Mt*DTBS, PDB entry 6czd |
|---|---|---|---|
| Data collection | | | |
| Diffraction source | Australian Synchrotron beamline MX1 | Australian Synchrotron beamline MX1 | Australian Synchrotron beamline MX1 |
| Wavelength (Å) | 0.9537 | 0.9537 | 0.9537 |
| Temperature (K) | 100 | 100 | 100 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| *a*, *b*, *c* (Å) | 55.33, 105.29, 153.76 | 56.680, 105.70, 153.26 | 54.77, 105.75, 155.20 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution range (Å) | 52.650–2.500 (2.600–2.500) | 56.680–2.500 (2.600–2.500) | 51.650–2.400 (2.490–2.400) |
| Total No. of reflections | 394104 (39657) | 428391 (46135) | 367683 (33793) |
| No. of unique reflections | 31792 (3444) | 32594 (3577) | 33558 (2787) |
| Completeness (%) | 99.500 (97.500) | 99.700 (98.700) | 93.400 (78.800) |
| Multiplicity | 12.400 (11.500) | 13.100 (12.900) | 10.900 (11.600) |
| $\langle I/\sigma(I)\rangle$ | 7.100 (1.4) | 7.100 (0.9) | 18.500 (7.7) |
| $R_{\rm r.i.m.}$ | 0.367 (2.293) | 0.449 (3.796) | 0.101 (0.291) |
| $R_{\rm p.i.m.}$ | 0.102 (0.653) | 0.122 (1.041) | 0.029 (0.079) |
| Overall *B* factor from Wilson plot (Å$^2$) | 36.700† | 41.770† | 28.200† |
| Refinement | | | |
| $\sigma$ Cutoff | $F > 0.310\sigma(F)$ | $F > 0.320\sigma(F)$ | $F > 0.310\sigma(F)$ |
| No. of reflections, working set | 30053 (2584) | 30798 (2470) | 31802 (2194) |
| No. of reflections, test set | 1603 (154) | 1639 (140) | 1742 (130) |
| Final $R_{\rm cryst}$ | 0.210 (0.2852) | 0.217 (0.3519) | 0.189 (0.1956) |
| Final $R_{\rm free}$ | 0.285 (0.3614) | 0.288 (0.4109) | 0.259 (0.3263) |
| No. of non-H atoms | | | |
| Protein | 6351 | 6351 | 6306 |
| Ligand | 120 | 120 | 112 |
| Solvent | 377 | 319 | 596 |
| Total | 6848 | 6790 | 7014 |
| R.m.s. deviations | | | |
| Bonds (Å) | 0.009 | 0.009 | 0.007 |
| Angles (°) | 1.248 | 1.301 | 0.999 |
| Average *B* factors (Å$^2$) | | | |
| Protein | 41.28 | 54.56 | 31.23 |
| Ligand | 35.13 | 30.84 | 33.86 |
| Ramachandran plot | | | |
| Most favoured (%) | 98.22 | 97.78 | 97.56 |
| Allowed (%) | 1.44 | 1.89 | 1.89 |

† Anomalies in the Wilson plot arose owing to ice rings.

that were stable in solution for more than 2 d, depending upon the ligand used. This method served as a template for obtaining previously intractable co-complexes involving *Mt*DTBS.

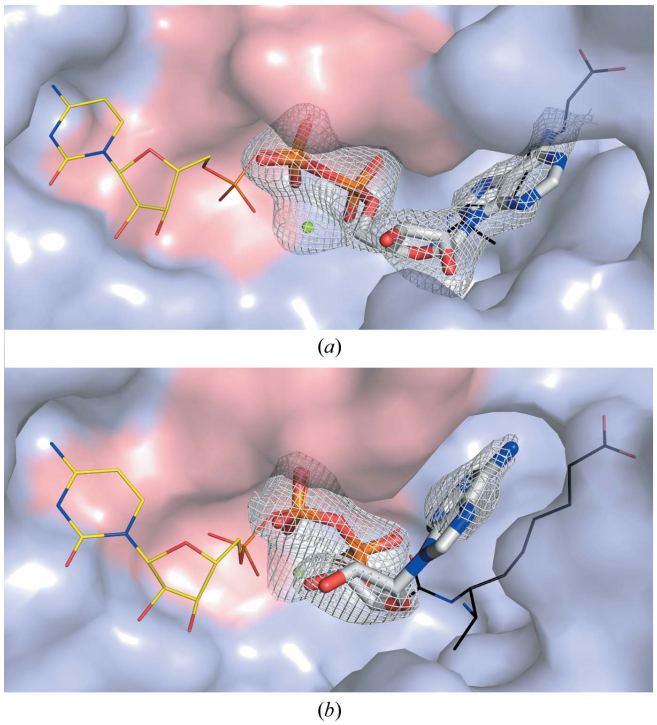
### 3.4. Proof of concept: precipitant–ligand exchange with CTP reveals the expected binding mode

To demonstrate that structures obtained using precipitant–ligand exchange are biologically relevant, crystals of *Mt*DTBS were grown in either the sodium citrate or ammonium sulfate crystallization conditions and then transferred into an exchange solution containing 100 m*M* CTP in place of the precipitant (Exchange Solutions 3 and 4, respectively; §2). Photographs of crystals in the transfer process are provided in Supplementary Fig. S1. After a short exchange (10–30 min), the *Mt*DTBS crystals were flash-cooled and subjected to X-ray data collection at the Australian Synchrotron. Both data sets were solved to 2.5 Å resolution by molecular replacement (PDB entries 6e05 and 6e06; crystallographic data are listed in Table 2). Consistent with previous reports, *Mt*DTBS was solved in space group $P2_12_12_1$, with four monomers in the asymmetric unit corresponding to two biological dimers. The monomer possessed the same $\alpha$-helical/$\beta$-sheet fold as previously reported, with the P-loop present between $\beta1$ and $\alpha1$ and the cytidine binding site formed by the coils connecting $\beta7$ to $3_{10}2$ and $\beta6$ to $\alpha7$ (Fig. 1). In both structures, CTP was modelled into well defined electron density present in the active site (Supplementary Figs. S2*a* and S2*b*), adopting an identical binding pose to our previous structure (PDB entry 4wop; Fig. 2). These data demonstrate that the complexes acquired by precipitant–ligand exchange were consistent with those obtained by conventional X-ray crystallographic techniques such as co-crystallization.

### 3.5. A crystal structure of ADP–*Mt*DTBS revealed an unexpected binding mode

Attempts to obtain a structure of *Mt*DTBS in complex with the inhibitor ADP using traditional crystal soaking were unsuccessful. We therefore turned to the precipitant–ligand exchange method, whereby *Mt*DTBS crystals were first grown in the ammonium sulfate condition and then transferred into Exchange Solution 5 containing 70 m*M* ADP in place of the ammonium sulfate. Under these conditions crystals were stable for several hours but dissolved overnight, encouraging the use of shorter soak times. The crystals were subjected to an hour-long exchange before flash-cooling and X-ray data collection at the Australian Synchrotron. The resulting data set was solved to a resolution of 2.4 Å by means of molecular replacement (PDB entry 6czd; crystallographic data are given in Table 2). Consistent with previous reports, *Mt*DTBS was solved in the same space group $P2_12_12_1$, with the overall fold and asymmetric unit as described above.

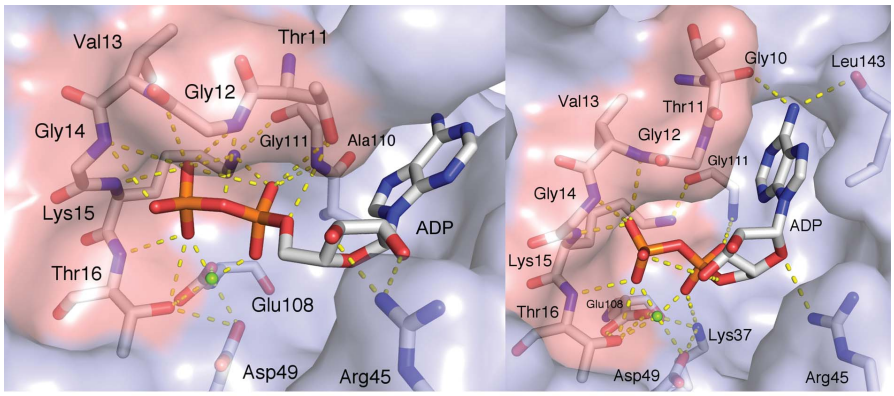
ADP was modelled into well defined electron density present in the active site (Figs. 4*a* and 4*b*; Supplementary Figs. S2*c* and S2*d*). In validation of the technique, no precipitant molecules were observed in the P-loop. Instead, the polyphosphate moiety of ADP was bound at this site. ADP bound in two alternate conformations, both of which extended in the reverse direction compared with the prototypical CTP conformation. These two conformations were differentiated by the adenosine moiety binding 'within' (*anti* conformation, present in chains *A*, *B* and *D*; Figs. 4*a* and 5*a*) or 'adjacent' to (*syn* conformation, present in chain *C*; Figs. 4*b* and 5*b*) the DAPA binding site. In both conformations the ADP $\beta$-phosphate bound at the same position as that observed for CTP (Salaemae *et al.*, 2015), in a highly coordinated network of hydrogen bonds and electrostatic interactions with the P-loop, involving the backbone amides of Val13 (3.2–3.42 Å), Gly14 (2.67–3.15 Å), Lys15 (2.47–3.05 Å) and Thr16 (3.04–3.1 Å), as well as the side chains of Lys15 (2.55–2.65 Å) and Thr16 (2.91–3.07 Å) (Figs. 5*a* and 5*b*). The ADP $\alpha$-phosphate bound at the CTP $\gamma$-phosphate position in both conformations; however, the protein interactions differed slightly. In the 'within' conformation, the $\alpha$-phosphate formed hydrogen-bonding and electrostatic interactions with the backbone of Gly111 (2.89–3.02 Å) and the side chain of Lys15 (2.39–2.48 Å) (Fig. 5*a*). In the 'adjacent' conformation, the $\alpha$-phosphate also interacted with the backbone of Gly111 (3.28 Å) but formed an interaction with the side chain of



(*a*)

(*b*)

**Figure 4**
Alternate ADP binding modes. ADP (grey sticks) and a magnesium ion (green sphere) were modelled into well defined electron density (white; polder map 3$\sigma$; Liebschner *et al.*, 2017) in the active site of *Mt*DTBS (light blue surface). Superposition with CTP (yellow lines; PDB entry 4wop) and DAPA carbamate (black lines; PDB entry 3fmf; Dey *et al.*, 2010) revealed that ADP clearly bound in the reverse direction to CTP; however, the $\beta$-phosphates for ADP and CTP bound in the same location at the P-loop (salmon surface). ADP exhibited two different conformations, (*a*) within and (*b*) adjacent to the DAPA pocket, with the latter accompanied by rearrangement of the P-loop.




(*a*) (*b*)

**Figure 5**
Polar intermolecular interactions between ADP (grey sticks) and *Mt*DTBS (blue surface, grey sticks). A highly coordinated network of hydrogen bonds and electrostatic interactions forms between the P-loop (salmon surface), the $\alpha$- and $\beta$-phosphate groups of ADP and the magnesium ion (green sphere). The magnesium ion exhibits six-coordination with the ADP phosphates, Thr16, Asp49, Glu108 and a water molecule (water molecules are not shown for clarity). Intermolecular interactions differ dependent on the adenosine conformation. (*a*) When bound within the DAPA pocket, the adenine moiety did not form polar interactions. (*b*) When bound adjacent to the DAPA pocket, Thr11 moved away from ADP, allowing polar contacts between the adenosine moiety and the carbonyl backbones of Gly10 and Leu143.

Lys37 (3.15 Å) (Fig. 4*b*). Both of the phosphates also bound to a magnesium present in the active site, which coordinated to the side chains of Thr16, Asp49 and Glu108, as well as a water molecule (all 1.83–2.44 Å; Figs. 5*a* and 5*b*). The ribose ring was hydrogen bonded to the side chain of Arg45 in both conformations; however, the 'within' conformation interacted through the 2′- and 3′-hydroxyl groups (2.55–3.06 Å; Fig. 5*a*), whereas it was flipped in the 'adjacent' conformation, interacting through O4′ (3.06 Å; Fig. 5*b*). When bound within the DAPA pocket, the adenine did not make polar contacts with the enzyme (Fig. 4*a*). Instead, the adenine was bound to water molecules and encased by hydrophobic contacts with Thr11, Met72, Ala73, Ala110, Val115 and Leu143. In this conformation, the adenine did not extend as deep into the binding site as the native substrate DAPA (Fig. 4*a*). When the adenine moiety bound adjacent to the DAPA pocket, rearrangement of the P-loop saw Thr11 move away from the ADP, allowing two hydrogen bonds to be made between the adenine N6 and the carbonyl backbone atoms of Gly10 (3.19 Å) and Leu143 (3.29 Å) (Fig. 5*b*).

## 4. Discussion

In this report, we name and describe precipitant–ligand exchange, a methodology to obtain structural data when crystallographic precipitants compete for binding with a ligand. This technique was validated by obtaining two crystal structures of *Mt*DTBS in complex with CTP. Precipitant–ligand exchange revealed the same binding mode for CTP as that obtained from other crystallographic techniques. We then demonstrated the practical utilization of this technique by obtaining crystallographic data for the previously intractable complex between ADP and *Mt*DTBS. We anticipate that this precipitant–ligand exchange approach may be widely applicable to ligand–protein interaction crystallography. However, several key points should be considered. This technique requires crystals that are amenable to soaking, a ligand of appropriate size for diffusion through solvent channels, and solvent-accessibility of the protein binding pocket. Under these conditions, ligand occupation may be affected by the following factors: the concentration and solubility of the ligand, the duration of interaction and, as observed here, the properties of competing precipitants. The affinity of the protein–ligand interaction is also an important consideration in the experimental design. For higher affinity protein–ligand interactions, such as CTP–*Mt*DTBS ($IC_{50}$ = 54.4 µ*M*; Salaemae *et al.*, 2015), precipitant occupation was easily overcome with a moderate concentration of ligand (CTP was co-crystallized at 1 m*M*; Salaemae *et al.*, 2015). However, the weaker affinity ADP interaction ($IC_{50}$ = 248 µ*M*; Salaemae, 2015) was outcompeted by the precipitant. Thus, the aim of this methodology was to achieve the optimal concentrations of both precipitant and ligand to result in distinct and observable electron density solely for the ligand after the exchange. Furthermore, ligand concentration and soaking duration were optimized such that crystal degradation was minimal, whilst still allowing high-occupancy ligand binding to occur. In order to aid other researchers attempting this technique, we have provided a flow diagram for decision making (Fig. 6).

To ensure that ligand density was unobscured, the precipitant was completely replaced with the ligand of interest. While crystal degradation did occur, shorter soak times were found to be sufficient for both diffraction quality and ligand density (*e.g.* 30–60 min). Some protein crystals may be less stable, and optimization of the ratio of ligand to precipitant may be required. This should be performed by solving several data sets at different ratios and soak times and optimizing each based on both diffraction quality and ligand electron density. In such cases, avoiding crystal transfer may also be preferred, and dilution of the precipitant can be performed directly in the drop. For example, adding 6 µl of a precipitant-free high ligand-concentration solution to a 2 µl drop would reduce the precipitant concentration by 75%. However, care must be taken to delineate between the electron density of the precipitant and ligand during interpretation.

ADP was observed to bind in two alternate conformations in the vicinity of the DAPA pocket, a mode distinct from that of CTP. We have previously shown that ADP ($IC_{50}$ = 248 µ*M*; Salaemae, 2015) and ATP ($IC_{50}$ = 349 µ*M*; Salaemae *et al.*, 2015) displace a competitive ATP fluorophore with comparable affinities. A modest increase in affinity could be attributed to the dual binding mode of ADP; however, comparative

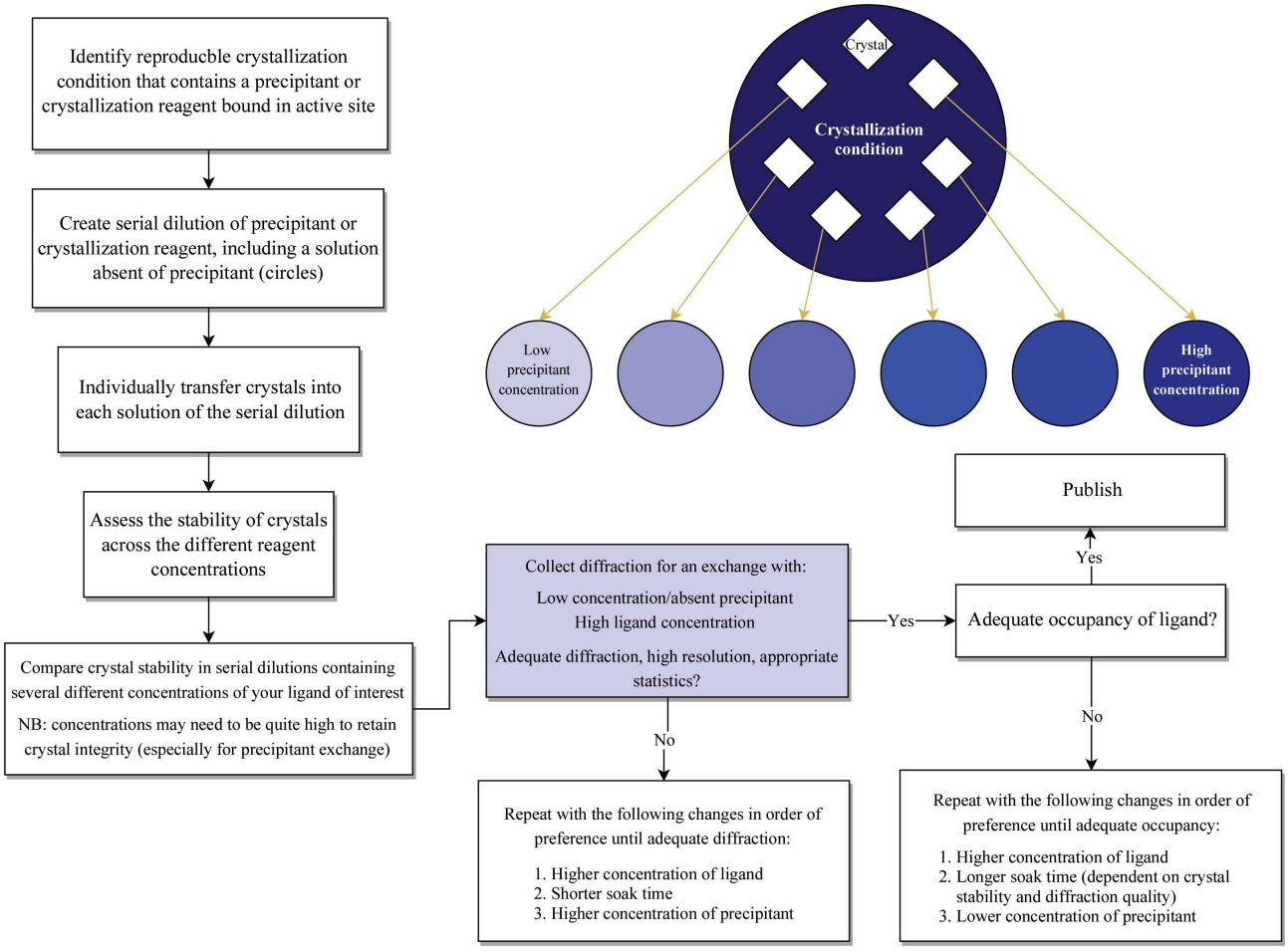



**Figure 6**
Flow diagram of the decision-making process during precipitant–ligand exchange.

structural investigation into the binding mode of ATP is ongoing. Interestingly, the $\beta$-phosphate of ADP in both conformations shared the same binding site as the $\beta$-phosphate of CTP within the P-loop. This was also the same site as observed for the sulfate precipitant, suggesting that this is a high-affinity binding site for anionic compounds. These structural data provide the first evidence for a molecule other than DAPA carbamate, or the closely related molecule 7-keto-8-aminopelargonic acid (KAPA), bound within the DAPA pocket. This revealed that the DAPA pocket can also accommodate larger and more varied compounds than the simple elongated acyl chains of DAPA and KAPA, a finding that will aid drug discovery. Finally, this structure indicates that inhibition by ADP may occur through competition with either substrate, a finding that invites further experimental investigation. The value of this structure lies in its contribution to mapping the binding characteristics of the active site of *Mt*DTBS. This will inform future antitubercular drug-discovery efforts targeting *Mt*DTBS activity and biotin biosynthesis, assisting in the design of inhibitors that bind with high specificity to the enzyme.

This technique may be of use in fragment-based drug discovery, where hits generally have weak affinity ($K_d$ > 500 µ*M*; Baker, 2013). If a fragment has been shown to bind to a protein using affinity and/or competitive inhibition assays, but only precipitant electron density is observed in the active site of the crystal structure, this technique can be used to exchange the precipitant, leading to fragment electron density. With weak ligands, it may not be possible to reduce the precipitant concentration to the required levels while maintaining crystal stability. This may lead to a scenario in which a certain proportion of protein molecules in the crystal are bound to precipitant, with the remainder bound to ligand. In this case, both precipitant- and ligand-bound states could be modelled (Pearce, Krojer & von Delft, 2017), or it may be useful to combine this technique with a program such as *PanDDA* (Pearce, Krojer, Bradley *et al.*, 2017) or polder maps (Liebschner *et al.*, 2017) to aid in discerning fragment density.

## 5. Conclusions

Here, we describe a crystallographic method for simultaneously removing and replacing active-site-bound precipitant molecules with a ligand of interest: a process that we have termed precipitant–ligand exchange. This technique was used to solve two CTP–*Mt*DTBS structures from independent

crystal-growth conditions. In validation of precipitant–ligand exchange, the results were consistent with the published literature. Subsequently, we solved a previously intractable structure of ADP bound to *Mt*DTBS. This structure revealed an alternate binding mode to that previously reported for CTP, providing new insight into the biology of the enzyme, as well as informing antitubercular inhibitor-design efforts. Given the prevalence of crystallographic reagents modelled in entries in the PDB, we believe that this method will be broadly applicable and will benefit structure-based interaction-analysis projects that utilize crystallography.

## Acknowledgements

This research was undertaken on the MX1 beamline at the Australian Synchrotron, which is part of ANSTO.

## Funding information

AT is the recipient of a University of Adelaide Postgraduate Scholarship. This work was supported by funding from the Channel 7 Children's Research Foundation (Project 181614).

## References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Baker, M. (2013). *Nature Rev. Drug Discov.* **12**, 5–7.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Ciccone, L., Vera, L., Tepshi, L., Rosalia, L., Rossello, A. & Stura, E. A. (2015). *Biotechnol. Rep.* **7**, 120–127.
Dai, R., Geders, T. W., Liu, F., Park, S. W., Schnappinger, D., Aldrich, C. C. & Finzel, B. C. (2015). *J. Med. Chem.* **58**, 5208–5217.
Dai, R., Wilson, D. J., Geders, T. W., Aldrich, C. C. & Finzel, B. C. (2014). *Chembiochem*, **15**, 575–586.
Dey, S., Lane, J. M., Lee, R. E., Rubin, E. J. & Sacchettini, J. C. (2010). *Biochemistry*, **49**, 6746–6760.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Huang, W., Jia, J., Gibson, K. J., Taylor, W. S., Rendina, A. R., Schneider, G. & Lindqvist, Y. (1995). *Biochemistry*, **34**, 10985–10995.
Käck, H., Gibson, K. J., Lindqvist, Y. & Schneider, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5495–5500.
Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
Karplus, P. A. & Diederichs, K. (2015). *Curr. Opin. Struct. Biol.* **34**, 60–68.
Kontopidis, G., Andrews, M. J. I., McInnes, C., Cowan, A., Powers, H., Innes, L., Plater, A., Griffiths, G., Paterson, D., Zheleva, D. I., Lane, D. P., Green, S., Walkinshaw, M. D. & Fischer, P. M. (2003). *Structure*, **11**, 1537–1546.
Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C. & Adams, P. D. (2017). *Acta Cryst.* D**73**, 148–157.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
Öster, L., Tapani, S., Xue, Y. & Käck, H. (2015). *Drug Discov. Today*, **20**, 1104–1111.
Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M. & von Delft, F. (2017). *Nature Commun.* **8**, 15123.
Pearce, N. M., Krojer, T. & von Delft, F. (2017). *Acta Cryst.* D**73**, 256–266.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Salaemae, W. (2015). PhD thesis, pp. 70–74. University of Adelaide, Australia.
Salaemae, W., Azhar, A., Booker, G. W. & Polyak, S. W. (2011). *Protein Cell*, **2**, 691–695.
Salaemae, W., Booker, G. W. & Polyak, S. W. (2016). *Microbiol Spectrum*, **4**, https://doi.org/10.1128/microbiolspec.VMBF-0008-2015.
Salaemae, W., Yap, M. Y., Wegener, K. L., Booker, G. W., Wilce, M. C. J. & Polyak, S. W. (2015). *Tuberculosis*, **95**, 259–266.
Thompson, A. P., Sternicki, L. M., Wegener, K. L., Lu, W., Zuo, L., Booker, G. W., Polyak, S. W. & Li, Y. (2016). *Jiangsu J. Prevent. Med.* **27**, 257–261.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Woong Park, S., Klotzsche, M., Wilson, D. J., Boshoff, H. I., Eoh, H., Manjunatha, U., Blumenthal, A., Rhee, K., Barry, C. E., Aldrich, C. C., Ehrt, S. & Schnappinger, D. (2011). *PLoS Pathog.* **7**, e1002264.

## Supporting information

# Precipitant–ligand exchange technique reveals the ADP binding mode in Mycobacterium tuberculosis dethiobiotin synthetase

**Authors**

Andrew P. Thompson,[a] Kate L. Wegener,[b] Grant W. Booker,[a] Steven W. Polyak[a] and John B. Bruning[b]*

**Affiliations**

[a]Molecular and Biomedical Science, The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia,

[b]Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia.

*Correspondence e-mail: john.bruning@adelaide.edu.au

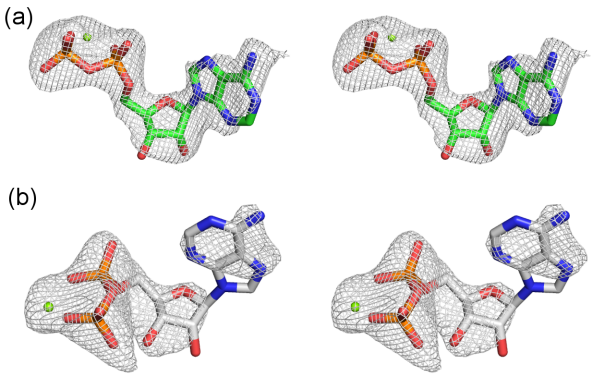



**Figure S1:** Walleye Polder map (3σ) of ADP bound in two different conformations: (a) the 'within' DAPA pocket conformation (b) the 'adjacent' to DAPA pocket conformation.

# Chapter 3

Publication: *Mycobacterium tuberculosis* dethiobiotin synthetase facilitates nucleoside triphosphate promiscuity through alternate binding modes

## 2.1 Statement of Authorship

| | |
|---|---|
| Title of Paper | Precipitant-ligand exchange technique reveals ADP binding mode in *Mt*DTBS |
| Publication Status | Published |
| Publication Details | Thompson, A. P., Wegener, K. L., Booker, G. W., Polyak, S. W. & John B. Bruning (2018). *Acta Cryst. D***74**. https://doi.org/10.1107/S2059798318010136. |

## Principal Author

| | | | |
|---|---|---|---|
| Name of Principal Author (Candidate) | Andrew P. Thompson | | |
| Contribution to the Paper | Performed all experimental work and analysis, including expression of *Mt*DTBS enzyme, crystallographic experiments including structure solving and interpretation. Prepared figures and manuscript. | | |
| Overall percentage (%) | 80% | | |
| Certification: | This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper. | | |
| Signature | | Date | 20/11/18 |

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

| | | | |
|---|---|---|---|
| Name of Co-Author | Kate L. Wegener | | |
| Contribution to the Paper | Assisted in experimental design, manuscript preparation and proof-reading. | | |
| Signature | | Date | 5/12/18 |

| Name of Co-Author | Grant W. Booker | | |
|---|---|---|---|
| Contribution to the Paper | Assisted manuscript preparation and proof-reading | | |
| Signature | | Date | 20/11/2018 |

| Name of Co-Author | Steven W. Polyak | | |
|---|---|---|---|
| Contribution to the Paper | Conceived *Mt*DTBS project. Assisted in experimental design, manuscript preparation and proof-reading. | | |
| Signature | | Date | 23/11/2018 |

| Name of Co-Author | John B. Bruning | | |
|---|---|---|---|
| Contribution to the Paper | Conceived and guided crystallographic experiments, helped with data processing and interpretation, assisted manuscript preparation and proof-reading. Corresponding author. | | |
| Signature | | Date | 20/11/2018 |

STRUCTURAL BIOLOGY
ISSN 2059-7983

# Precipitant–ligand exchange technique reveals the ADP binding mode in *Mycobacterium tuberculosis* dethiobiotin synthetase

**Andrew P. Thompson,[a] Kate L. Wegener,[b] Grant W. Booker,[a] Steven W. Polyak[a] and John B. Bruning[b]***

[a]Molecular and Biomedical Science, The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia, and [b]Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia. *Correspondence e-mail: john.bruning@adelaide.edu.au

Received 22 May 2018
Accepted 13 July 2018

Edited by B. Kobe, University of Queensland, Australia

**Keywords:** dethiobiotin synthetase; *Mycobacterium tuberculosis*; crystallography; precipitant; nucleoside polyphosphate; ligand; exchange; soak.

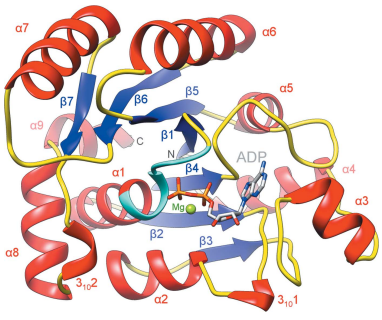
**PDB references**: ADP–*Mt*DTBS, 6czd; CTP–*Mt*DTBS, 6e05; 6e06

**Supporting information**: this article has supporting information at journals.iucr.org/d

Dethiobiotin synthetase from *Mycobacterium tuberculosis* (*Mt*DTBS) is a promising antituberculosis drug target. Small-molecule inhibitors that target *Mt*DTBS provide a route towards new therapeutics for the treatment of antibiotic-resistant tuberculosis. Adenosine diphosphate (ADP) is an inhibitor of *Mt*DTBS; however, structural studies into its mechanism of inhibition have been unsuccessful owing to competitive binding to the enzyme by crystallographic precipitants such as citrate and sulfate. Here, a crystallographic technique termed precipitant–ligand exchange has been developed to exchange protein-bound precipitants with ligands of interest. Proof of concept for the exchange method was demonstrated using cytidine triphosphate (CTP), which adopted the same binding mechanism as that obtained with traditional crystal-soaking techniques. Precipitant–ligand exchange also yielded the previously intractable structure of *Mt*DTBS in complex with ADP solved to 2.4 Å resolution. This result demonstrates the utility of precipitant–ligand exchange, which may be widely applicable to protein crystallography.

## 1. Introduction

Biotin biosynthesis has been identified as a promising pathway for the development of new antituberculosis chemotherapies required to combat the rise of multidrug-resistant infections (Salaemae *et al.*, 2011, 2016; Thompson *et al.*, 2016; Dai *et al.*, 2014, 2015; Dey *et al.*, 2010; Woong Park *et al.*, 2011). The penultimate reaction in the pathway is the energy-dependent conversion of 7,8-diaminopelargonic acid (DAPA) to dethiobiotin (DTB), catalysed by dethiobiotin synthetase (DTBS). Small molecules that target DTBS represent valuable starting points for antitubercular drug discovery. We have previously demonstrated that ADP is an inhibitor of *Mycobacterium tuberculosis* DTBS (*Mt*DTBS; $K_i$ = 0.14 m*M* with respect to ATP) and binds with comparable affinity to ATP (ADP $IC_{50}$ = 248 µ*M*, ATP $IC_{50}$ = 349 µ*M*) in a competitive displacement assay (Salaemae, 2015). Thus, ADP provides a potential starting point for the chemical development of nucleoside analogues to treat antibiotic-resistant tuberculosis. The crystal structure of *Mt*DTBS in complex with cytidine triphosphate (CTP; PDB entry 4wop) revealed the binding mode for this ligand (Salaemae *et al.*, 2015). The cytosine moiety bound in a pocket formed by the coils connecting $\beta$7 to $3_{10}$2 and $\beta$6 to $\alpha$7 (Figs. 1 and 2), and the polyphosphate group bound at the phosphate-binding loop (P-loop, Walker A motif; depicted in Fig. 1). However, efforts to crystallize *Mt*DTBS in complex with other nucleosides such as ADP have been unsuccessful



(Salaemae *et al.*, 2015; Dey *et al.*, 2010) despite similar structures having been obtained with homologues from other bacteria (Huang *et al.*, 1995; Käck *et al.*, 1998). In our repeated attempts to solve the ADP–*Mt*DTBS crystal structure, electron density consistent with a sulfate ion bound at the P-loop was observed, resulting from the high concentrations of ammonium sulfate in the crystallization condition (1.2–1.7 *M* ammonium sulfate, 0.1 *M* Tris pH 8, 10–15% glycerol; adapted from Salaemae *et al.*, 2015). Concordant with this, a sulfate ion was observed instead of CTP in one active site of the CTP–*Mt*DTBS crystal structure (Salaemae *et al.*, 2015). We propose that precipitants that are present at high concentrations in the crystallization condition directly compete with ADP for binding to the P-loop. Consequently, we sought to overcome precipitant occupation of the active site in order to investigate *Mt*DTBS in complex with ADP.

Crystallographic reagents commonly bind to distinct regions of protein molecules. Table 1 depicts the prevalence of common crystallographic reagents modelled in structures deposited as Protein Data Bank (PDB) entries. Of the total submissions in the PDB (139 357 at the time of writing), sulfate is present in 12.4% of submissions and glycerol (a common cryoprotectant) is present in 10.5%. These two molecules are the most commonly modelled free ligands in the PDB, indicating that this is a common occurrence in protein crystallography. Efforts to avoid crystallographic reagents may require either a new crystallization condition or the use of a homologous protein. Alternatively, techniques to remove these compounds by crystal manipulation and soaking have also been reported. Examples of such approaches in the literature include crystal washing to remove dioxane, back-soaking to remove substrates and soaking with homologous

**Table 1**
The prevalence of selected crystallographic reagents modelled as free ligands in Protein Data Bank entries at the time of writing (139 357 total structures).

| Crystallographic molecule | No. of entries | Percentage of total submissions |
|---|---|---|
| Sulfate ion | 17338 | 12.4 |
| Glycerol | 14594 | 10.5 |
| Ethane-1,2-diol (EDO) | 8370 | 6.0 |
| Phosphate ion | 5046 | 3.6 |
| Acetate ion | 4334 | 3.1 |
| Di(hydroxyethyl)ether (nonpolymer PEG) | 2444 | 1.8 |
| Citrate | 409 | 0.3 |

precipitants (for example, replacing citrate with malonate) (reviewed in Öster *et al.*, 2015). Ligand-exchange methods have previously been utilized to substitute co-crystallized peptides (Kontopidis *et al.*, 2003) and degraded substrates (Ciccone *et al.*, 2015) with ligands. In the current study, this principle was applied to exchange active-site-bound crystallographic precipitants with ligands. This alternate methodology was required to ensure the stability of *Mt*DTBS crystals, which was reliant upon a high concentration of a suitable compound (either precipitant or ligand) being bound within the P-loop.

Here, we describe a soaking methodology to overcome the occlusion of ligands of interest by crystallographic precipitant molecules in the active site. Having ascertained that alternate crystal-growth conditions also resulted in precipitants bound to the P-loop, we assessed the crystal integrity in varying concentrations of the precipitants sulfate and citrate, and ultimately replaced the precipitant with high concentrations of the ligand of interest. As proof of concept, two structures of *Mt*DTBS in complex with CTP were solved using precipitant–ligand exchange with crystals from two independent growth conditions. In both structures, CTP adopted the same binding pose as demonstrated by our previous crystal structure (PDB entry 4wop). Subsequently, a crystal structure of the previously intractable complex with the inhibitor ADP was obtained. We suggest that our precipitant–ligand exchange technique may be broadly applicable to crystallography of protein–ligand complexes and may be especially important for obtaining crystals of weak-affinity complexes such as those studied in fragment-based drug discovery.

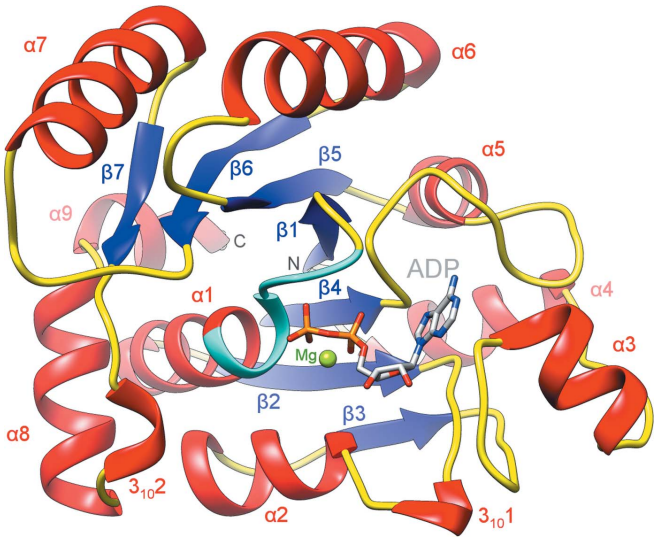



**Figure 1**
The ADP-bound monomer of *Mt*DTBS (chain *A*; PDB entry 6czd). Secondary structure is depicted as ribbons: helices are coloured red, $\beta$-sheets blue and interconnecting coils yellow. ADP (grey sticks) bound distal to the cytidine-binding site (coils connecting $\beta7$ to $3_{10}2$ and $\beta6$ to $\alpha7$). The $\beta$-phosphate bound at the P-loop (cyan) coordinated to a magnesium ion (green sphere).

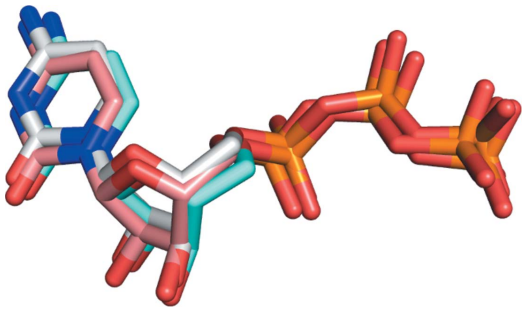


**Figure 2**
Superposition of CTP in our previously published structure (white sticks; PDB entry 4wop) and those solved by precipitant–ligand exchange (pink, sulfate-grown crystal; blue, citrate-grown crystal).

## 2. Materials and methods

### 2.1. Protein expression and purification

*Mt*DTBS was expressed and purified as described previously (Salaemae *et al.*, 2015), except that a HiPrep 26/60 Sephacryl S-300 HR column (GE Healthcare) was used as an additional purification step following IMAC chromatography. *Mt*DTBS was concentrated to 10 mg ml$^{-1}$ and stored in the following solution: 25 m*M* Tris pH 7.5, 30 m*M* NaCl, 1 m*M* EDTA, 1 m*M* dithiothreitol (DTT), 5%(*v*/*v*) glycerol.

### 2.2. Nucleoside polyphosphate preparation

Nucleoside polyphosphates (NpPs) were purchased as analytical grade powders (Sigma–Aldrich) and solubilized to 100–500 m*M* in Milli-Q $H_2O$ followed by adjustment to pH 7 with NaOH.

### 2.3. General crystallographic methods

All *Mt*DTBS crystals were grown *via* the hanging-drop method with a well volume of 500 µl and a drop ratio of 1:1 protein:well solution at 16°C unless otherwise stated. Sulfate ion-bound *Mt*DTBS crystals were grown in an ammonium sulfate crystallization condition consisting of 1.2–1.7 *M* ammonium sulfate, 0.1 *M* Tris pH 8, 10–15% glycerol. Citrate-bound *Mt*DTBS crystals were grown in a citrate crystallization solution consisting of 0.7 *M* sodium citrate, 0.1 *M* imidazole pH 7.0 with 20% glycerol.

### 2.4. Serial dilutions and ligand-exchange solutions

Serial Dilution Solution 1: a twofold serial dilution of citrate (0.35–0.011 *M*) was produced by mixing the citrate crystallization condition with an equal volume of precipitant-free solution (0.1 *M* imidazole pH 7.0 with 20% glycerol).

Serial Dilution Solution 2: a twofold serial dilution of citrate (0.35–0.011 *M*) in the presence of 5 m*M* ATP was produced by mixing the citrate crystallization condition with a solution containing ATP in place of citrate (10 m*M* ATP, 1 m*M* $MgCl_2$, 0.1 *M* imidazole pH 7.0, 20% glycerol).

Exchange Solution 1 consisted of 300 m*M* ATP, 60 m*M* $MgCl_2$, 0.022 *M* sodium citrate, 0.1 *M* imidazole pH 7.0, 20% glycerol.

Exchange Solution 2 consisted of 300 m*M* ATP, 30 m*M* $MgCl_2$, 0.1 *M* Tris pH 8, 10–20% glycerol. The well solution consisted of 0.1 *M* Tris pH 8, 10–20% glycerol.

Exchange Solution 3 consisted of 100 m*M* CTP, 10 m*M* $MgCl_2$, 0.1 *M* imidazole pH 7, 20% glycerol.

Exchange Solution 4 consisted of 100 m*M* CTP, 10 m*M* $MgCl_2$, 0.1 *M* Tris pH 8, 25% glycerol.

Exchange Solution 5 consisted of 70 m*M* ADP, 30 m*M* $MgCl_2$, 0.1 *M* Tris pH 8, 10–20% glycerol. The well solution consisted of 0.1 *M* Tris pH 8, 10–20% glycerol.

### 2.5. Data collection and processing

All crystals were flash-cooled in liquid nitrogen and subjected to X-ray diffraction on the MX1 beamline at the Australian Synchrotron, which is part of ANSTO (McPhillips *et al.*, 2002). 360 images were collected with 1° oscillation at 100 K and at a wavelength of 0.9537 Å. Data were indexed, scaled and merged using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (*CCP*4; Winn *et al.*, 2011). Resolution truncation was performed in accordance with previously described $CC_{1/2}$ cutoff values (Karplus & Diederichs, 2012, 2015). Phasing was performed with *Phaser-MR* (McCoy *et al.*, 2007) using an in-house search model. The structure was iteratively refined with manual modelling in *Coot* (Emsley & Cowtan, 2004) and refinement with *phenix.refine* (Adams *et al.*, 2010) until the *R* factors converged and significant difference density was accounted for. The structure was validated using *MolProbity* as part of the *PHENIX* package. Structure figures were generated with the *UCSF Chimera* package (Pettersen *et al.*, 2004) and *MacPyMOL* (Schrödinger). Coordinates and structure factors were deposited in the Protein Data Bank for ADP–*Mt*DTBS (PDB entry 6czd), CTP–*Mt*DTBS from crystals grown in ammonium sulfate (PDB entry 6e05) and in sodium citrate (PDB entry 6e06).

## 3. Results

### 3.1. Alternate crystal conditions consistently contained precipitants bound at the P-loop

In order to obtain crystals of nucleoside polyphosphates (NpPs) bound to *Mt*DTBS without a competing precipitant, alternate crystallization conditions were explored. To avoid the use of ammonium sulfate, *Mt*DTBS crystals were grown in a citrate crystallization solution adapted from Dey *et al.* (2010) (see §2). However, no NpP-complexed crystal structures were obtained *via* either co-crystallization or soaking. Instead, electron density resembling citrate was observed bound to the P-loop in all solved crystal structures (data not shown). Further crystallographic screening yielded hits in conditions containing high concentrations of sulfate, citrate or organic acid mixtures (Tacsimate), all of which are likely to bind at the same location. Therefore, a technique was developed to replace the precipitants with an NpP.

### 3.2. Reducing the precipitant concentration caused crystal instability

The stability of *Mt*DTBS crystals in decreasing concentrations of precipitant was assessed. *Mt*DTBS crystals were grown in the citrate crystallization condition before being transferred directly into each solution of a twofold serial dilution of citrate (0.35–0.011 *M* sodium citrate; Serial Dilution Solution 1; §2). As expected, crystal stability correlated with the precipitant concentration (Fig. 3). Crystals transferred into solutions containing less than 0.044 *M* citrate dissolved rapidly (<30 min), those in 0.088 *M* citrate dissolved overnight and those in 0.175 *M* citrate cracked severely. It was therefore hypothesized that crystal integrity was dependent on the presence of a suitable molecule, in this case the precipitant, bound to the P-loop of *Mt*DTBS.

### 3.3. The addition of ligand increased the stability of crystals

As NpP molecules also bind to the P-loop, it was proposed that these molecules would rescue crystal stability in the absence of precipitant. To test this, crystal stability was assessed in a twofold serial dilution series of citrate containing a constant concentration of 5 m*M* ATP (Serial Dilution Solution 2; §2). There was no appreciable improvement in crystal stability using this concentration of ligand (Table 2). It was hypothesized that the percentage of *Mt*DTBS molecules in the crystal bound to a ligand (precipitant or NpP) was not high enough to retain crystal integrity. Therefore, much higher concentrations of NpP were trialled, analogous to the high concentrations of precipitant required for crystal growth described previously. Crystals were transferred into a solution of 0.022 *M* citrate that was supplemented with 300 m*M* ATP (Exchange Solution 1; §2). The crystals of *Mt*DTBS showed improved stability, as they did not degrade overnight. Similarly, crystals grown in sulfate were stable when supplemented with ATP. This was consistent with the hypothesis that crystal integrity was dependent on the presence of a suitable ligand to occupy the P-loop.

To ensure that the ligand of interest would be unobscured by precipitant density in the resulting crystal structure, efforts were made to completely remove the protein-bound precipitant. *Mt*DTBS crystals were grown in the ammonium sulfate crystallization condition before being transferred into an 8 µl drop of Exchange Solution 2 containing 300 m*M* ATP. An 8 µl drop size was used to minimize dilution by any of the original buffer components in the transfer process. This resulted in crystals

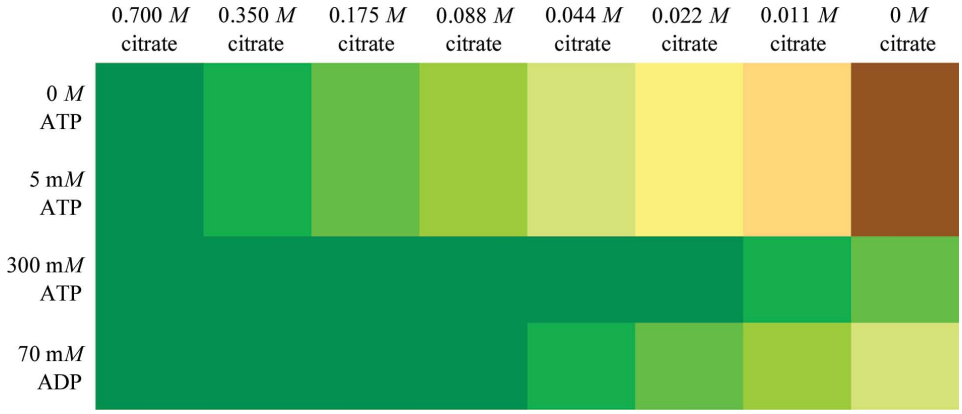



**Figure 3**
Qualitative crystal stability after transfer into solutions containing varying concentrations of citrate precipitant and ATP or ADP. Crystal stability is denoted from highest to lowest from green to brown.

**Table 2**
Data-collection and processing statistics for *Mt*DTBS crystal structures.

Values in parentheses are for the outer shell.

| | CTP–*Mt*DTBS (sulfate), PDB entry 6e05 | CTP–*Mt*DTBS (citrate), PDB entry 6e06 | ADP–*Mt*DTBS, PDB entry 6czd |
|---|---|---|---|
| Data collection | | | |
| Diffraction source | Australian Synchrotron beamline MX1 | Australian Synchrotron beamline MX1 | Australian Synchrotron beamline MX1 |
| Wavelength (Å) | 0.9537 | 0.9537 | 0.9537 |
| Temperature (K) | 100 | 100 | 100 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| *a*, *b*, *c* (Å) | 55.33, 105.29, 153.76 | 56.680, 105.70, 153.26 | 54.77, 105.75, 155.20 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution range (Å) | 52.650–2.500 (2.600–2.500) | 56.680–2.500 (2.600–2.500) | 51.650–2.400 (2.490–2.400) |
| Total No. of reflections | 394104 (39657) | 428391 (46135) | 367683 (33793) |
| No. of unique reflections | 31792 (3444) | 32594 (3577) | 33558 (2787) |
| Completeness (%) | 99.500 (97.500) | 99.700 (98.700) | 93.400 (78.800) |
| Multiplicity | 12.400 (11.500) | 13.100 (12.900) | 10.900 (11.600) |
| $\langle I/\sigma(I)\rangle$ | 7.100 (1.4) | 7.100 (0.9) | 18.500 (7.7) |
| $R_{\rm r.i.m.}$ | 0.367 (2.293) | 0.449 (3.796) | 0.101 (0.291) |
| $R_{\rm p.i.m.}$ | 0.102 (0.653) | 0.122 (1.041) | 0.029 (0.079) |
| Overall *B* factor from Wilson plot (Å$^2$) | 36.700† | 41.770† | 28.200† |
| Refinement | | | |
| $\sigma$ Cutoff | $F > 0.310\sigma(F)$ | $F > 0.320\sigma(F)$ | $F > 0.310\sigma(F)$ |
| No. of reflections, working set | 30053 (2584) | 30798 (2470) | 31802 (2194) |
| No. of reflections, test set | 1603 (154) | 1639 (140) | 1742 (130) |
| Final $R_{\rm cryst}$ | 0.210 (0.2852) | 0.217 (0.3519) | 0.189 (0.1956) |
| Final $R_{\rm free}$ | 0.285 (0.3614) | 0.288 (0.4109) | 0.259 (0.3263) |
| No. of non-H atoms | | | |
| Protein | 6351 | 6351 | 6306 |
| Ligand | 120 | 120 | 112 |
| Solvent | 377 | 319 | 596 |
| Total | 6848 | 6790 | 7014 |
| R.m.s. deviations | | | |
| Bonds (Å) | 0.009 | 0.009 | 0.007 |
| Angles (°) | 1.248 | 1.301 | 0.999 |
| Average *B* factors (Å$^2$) | | | |
| Protein | 41.28 | 54.56 | 31.23 |
| Ligand | 35.13 | 30.84 | 33.86 |
| Ramachandran plot | | | |
| Most favoured (%) | 98.22 | 97.78 | 97.56 |
| Allowed (%) | 1.44 | 1.89 | 1.89 |

† Anomalies in the Wilson plot arose owing to ice rings.

that were stable in solution for more than 2 d, depending upon the ligand used. This method served as a template for obtaining previously intractable co-complexes involving *Mt*DTBS.

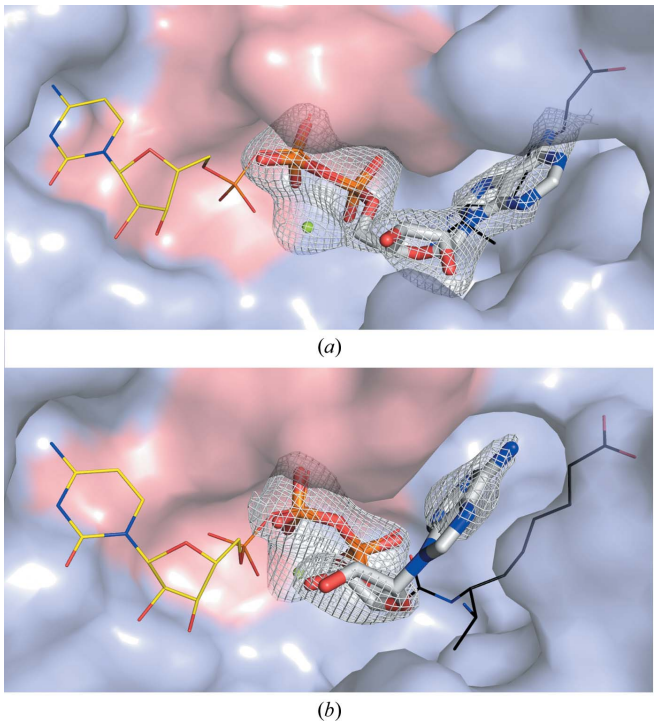
### 3.4. Proof of concept: precipitant–ligand exchange with CTP reveals the expected binding mode

To demonstrate that structures obtained using precipitant–ligand exchange are biologically relevant, crystals of *Mt*DTBS were grown in either the sodium citrate or ammonium sulfate crystallization conditions and then transferred into an exchange solution containing 100 m*M* CTP in place of the precipitant (Exchange Solutions 3 and 4, respectively; §2). Photographs of crystals in the transfer process are provided in Supplementary Fig. S1. After a short exchange (10–30 min), the *Mt*DTBS crystals were flash-cooled and subjected to X-ray data collection at the Australian Synchrotron. Both data sets were solved to 2.5 Å resolution by molecular replacement (PDB entries 6e05 and 6e06; crystallographic data are listed in Table 2). Consistent with previous reports, *Mt*DTBS was solved in space group $P2_12_12_1$, with four monomers in the asymmetric unit corresponding to two biological dimers. The monomer possessed the same $\alpha$-helical/$\beta$-sheet fold as previously reported, with the P-loop present between $\beta1$ and $\alpha1$ and the cytidine binding site formed by the coils connecting $\beta7$ to $3_{10}2$ and $\beta6$ to $\alpha7$ (Fig. 1). In both structures, CTP was modelled into well defined electron density present in the active site (Supplementary Figs. S2*a* and S2*b*), adopting an identical binding pose to our previous structure (PDB entry 4wop; Fig. 2). These data demonstrate that the complexes acquired by precipitant–ligand exchange were consistent with those obtained by conventional X-ray crystallographic techniques such as co-crystallization.

### 3.5. A crystal structure of ADP–*Mt*DTBS revealed an unexpected binding mode

Attempts to obtain a structure of *Mt*DTBS in complex with the inhibitor ADP using traditional crystal soaking were unsuccessful. We therefore turned to the precipitant–ligand exchange method, whereby *Mt*DTBS crystals were first grown in the ammonium sulfate condition and then transferred into Exchange Solution 5 containing 70 m*M* ADP in place of the ammonium sulfate. Under these conditions crystals were stable for several hours but dissolved overnight, encouraging the use of shorter soak times. The crystals were subjected to an hour-long exchange before flash-cooling and X-ray data collection at the Australian Synchrotron. The resulting data set was solved to a resolution of 2.4 Å by means of molecular replacement (PDB entry 6czd; crystallographic data are given in Table 2). Consistent with previous reports, *Mt*DTBS was solved in the same space group $P2_12_12_1$, with the overall fold and asymmetric unit as described above.

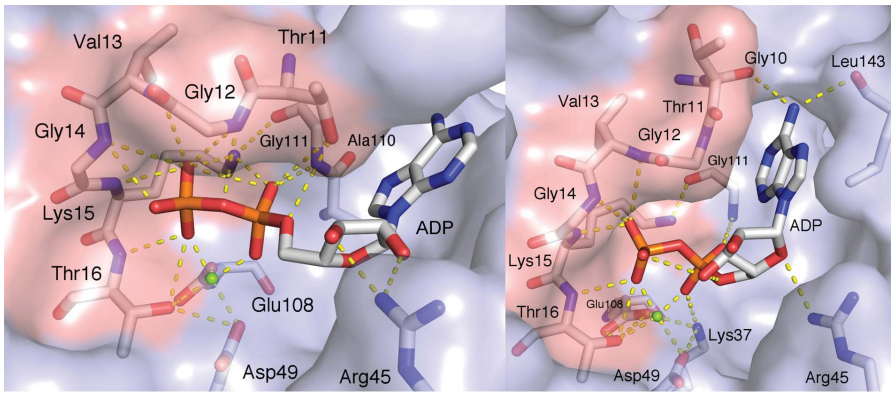
ADP was modelled into well defined electron density present in the active site (Figs. 4*a* and 4*b*; Supplementary Figs. S2*c* and S2*d*). In validation of the technique, no precipitant molecules were observed in the P-loop. Instead, the polyphosphate moiety of ADP was bound at this site. ADP bound in two alternate conformations, both of which extended in the reverse direction compared with the prototypical CTP conformation. These two conformations were differentiated by the adenosine moiety binding 'within' (*anti* conformation, present in chains *A*, *B* and *D*; Figs. 4*a* and 5*a*) or 'adjacent' to (*syn* conformation, present in chain *C*; Figs. 4*b* and 5*b*) the DAPA binding site. In both conformations the ADP $\beta$-phosphate bound at the same position as that observed for CTP (Salaemae *et al.*, 2015), in a highly coordinated network of hydrogen bonds and electrostatic interactions with the P-loop, involving the backbone amides of Val13 (3.2–3.42 Å), Gly14 (2.67–3.15 Å), Lys15 (2.47–3.05 Å) and Thr16 (3.04–3.1 Å), as well as the side chains of Lys15 (2.55–2.65 Å) and Thr16 (2.91–3.07 Å) (Figs. 5*a* and 5*b*). The ADP $\alpha$-phosphate bound at the CTP $\gamma$-phosphate position in both conformations; however, the protein interactions differed slightly. In the 'within' conformation, the $\alpha$-phosphate formed hydrogen-bonding and electrostatic interactions with the backbone of Gly111 (2.89–3.02 Å) and the side chain of Lys15 (2.39–2.48 Å) (Fig. 5*a*). In the 'adjacent' conformation, the $\alpha$-phosphate also interacted with the backbone of Gly111 (3.28 Å) but formed an interaction with the side chain of



(*a*)

(*b*)

**Figure 4**
Alternate ADP binding modes. ADP (grey sticks) and a magnesium ion (green sphere) were modelled into well defined electron density (white; polder map 3$\sigma$; Liebschner *et al.*, 2017) in the active site of *Mt*DTBS (light blue surface). Superposition with CTP (yellow lines; PDB entry 4wop) and DAPA carbamate (black lines; PDB entry 3fmf; Dey *et al.*, 2010) revealed that ADP clearly bound in the reverse direction to CTP; however, the $\beta$-phosphates for ADP and CTP bound in the same location at the P-loop (salmon surface). ADP exhibited two different conformations, (*a*) within and (*b*) adjacent to the DAPA pocket, with the latter accompanied by rearrangement of the P-loop.




(*a*) (*b*)

**Figure 5**
Polar intermolecular interactions between ADP (grey sticks) and *Mt*DTBS (blue surface, grey sticks). A highly coordinated network of hydrogen bonds and electrostatic interactions forms between the P-loop (salmon surface), the $\alpha$- and $\beta$-phosphate groups of ADP and the magnesium ion (green sphere). The magnesium ion exhibits six-coordination with the ADP phosphates, Thr16, Asp49, Glu108 and a water molecule (water molecules are not shown for clarity). Intermolecular interactions differ dependent on the adenosine conformation. (*a*) When bound within the DAPA pocket, the adenine moiety did not form polar interactions. (*b*) When bound adjacent to the DAPA pocket, Thr11 moved away from ADP, allowing polar contacts between the adenosine moiety and the carbonyl backbones of Gly10 and Leu143.

Lys37 (3.15 Å) (Fig. 4*b*). Both of the phosphates also bound to a magnesium present in the active site, which coordinated to the side chains of Thr16, Asp49 and Glu108, as well as a water molecule (all 1.83–2.44 Å; Figs. 5*a* and 5*b*). The ribose ring was hydrogen bonded to the side chain of Arg45 in both conformations; however, the 'within' conformation interacted through the 2′- and 3′-hydroxyl groups (2.55–3.06 Å; Fig. 5*a*), whereas it was flipped in the 'adjacent' conformation, interacting through O4′ (3.06 Å; Fig. 5*b*). When bound within the DAPA pocket, the adenine did not make polar contacts with the enzyme (Fig. 4*a*). Instead, the adenine was bound to water molecules and encased by hydrophobic contacts with Thr11, Met72, Ala73, Ala110, Val115 and Leu143. In this conformation, the adenine did not extend as deep into the binding site as the native substrate DAPA (Fig. 4*a*). When the adenine moiety bound adjacent to the DAPA pocket, rearrangement of the P-loop saw Thr11 move away from the ADP, allowing two hydrogen bonds to be made between the adenine N6 and the carbonyl backbone atoms of Gly10 (3.19 Å) and Leu143 (3.29 Å) (Fig. 5*b*).

## 4. Discussion

In this report, we name and describe precipitant–ligand exchange, a methodology to obtain structural data when crystallographic precipitants compete for binding with a ligand. This technique was validated by obtaining two crystal structures of *Mt*DTBS in complex with CTP. Precipitant–ligand exchange revealed the same binding mode for CTP as that obtained from other crystallographic techniques. We then demonstrated the practical utilization of this technique by obtaining crystallographic data for the previously intractable complex between ADP and *Mt*DTBS. We anticipate that this precipitant–ligand exchange approach may be widely applicable to ligand–protein interaction crystallography. However, several key points should be considered. This technique requires crystals that are amenable to soaking, a ligand of appropriate size for diffusion through solvent channels, and solvent-accessibility of the protein binding pocket. Under these conditions, ligand occupation may be affected by the following factors: the concentration and solubility of the ligand, the duration of interaction and, as observed here, the properties of competing precipitants. The affinity of the protein–ligand interaction is also an important consideration in the experimental design. For higher affinity protein–ligand interactions, such as CTP–*Mt*DTBS ($IC_{50}$ = 54.4 µ*M*; Salaemae *et al.*, 2015), precipitant occupation was easily overcome with a moderate concentration of ligand (CTP was co-crystallized at 1 m*M*; Salaemae *et al.*, 2015). However, the weaker affinity ADP interaction ($IC_{50}$ = 248 µ*M*; Salaemae, 2015) was outcompeted by the precipitant. Thus, the aim of this methodology was to achieve the optimal concentrations of both precipitant and ligand to result in distinct and observable electron density solely for the ligand after the exchange. Furthermore, ligand concentration and soaking duration were optimized such that crystal degradation was minimal, whilst still allowing high-occupancy ligand binding to occur. In order to aid other researchers attempting this technique, we have provided a flow diagram for decision making (Fig. 6).

To ensure that ligand density was unobscured, the precipitant was completely replaced with the ligand of interest. While crystal degradation did occur, shorter soak times were found to be sufficient for both diffraction quality and ligand density (*e.g.* 30–60 min). Some protein crystals may be less stable, and optimization of the ratio of ligand to precipitant may be required. This should be performed by solving several data sets at different ratios and soak times and optimizing each based on both diffraction quality and ligand electron density. In such cases, avoiding crystal transfer may also be preferred, and dilution of the precipitant can be performed directly in the drop. For example, adding 6 µl of a precipitant-free high ligand-concentration solution to a 2 µl drop would reduce the precipitant concentration by 75%. However, care must be taken to delineate between the electron density of the precipitant and ligand during interpretation.

ADP was observed to bind in two alternate conformations in the vicinity of the DAPA pocket, a mode distinct from that of CTP. We have previously shown that ADP ($IC_{50}$ = 248 µ*M*; Salaemae, 2015) and ATP ($IC_{50}$ = 349 µ*M*; Salaemae *et al.*, 2015) displace a competitive ATP fluorophore with comparable affinities. A modest increase in affinity could be attributed to the dual binding mode of ADP; however, comparative

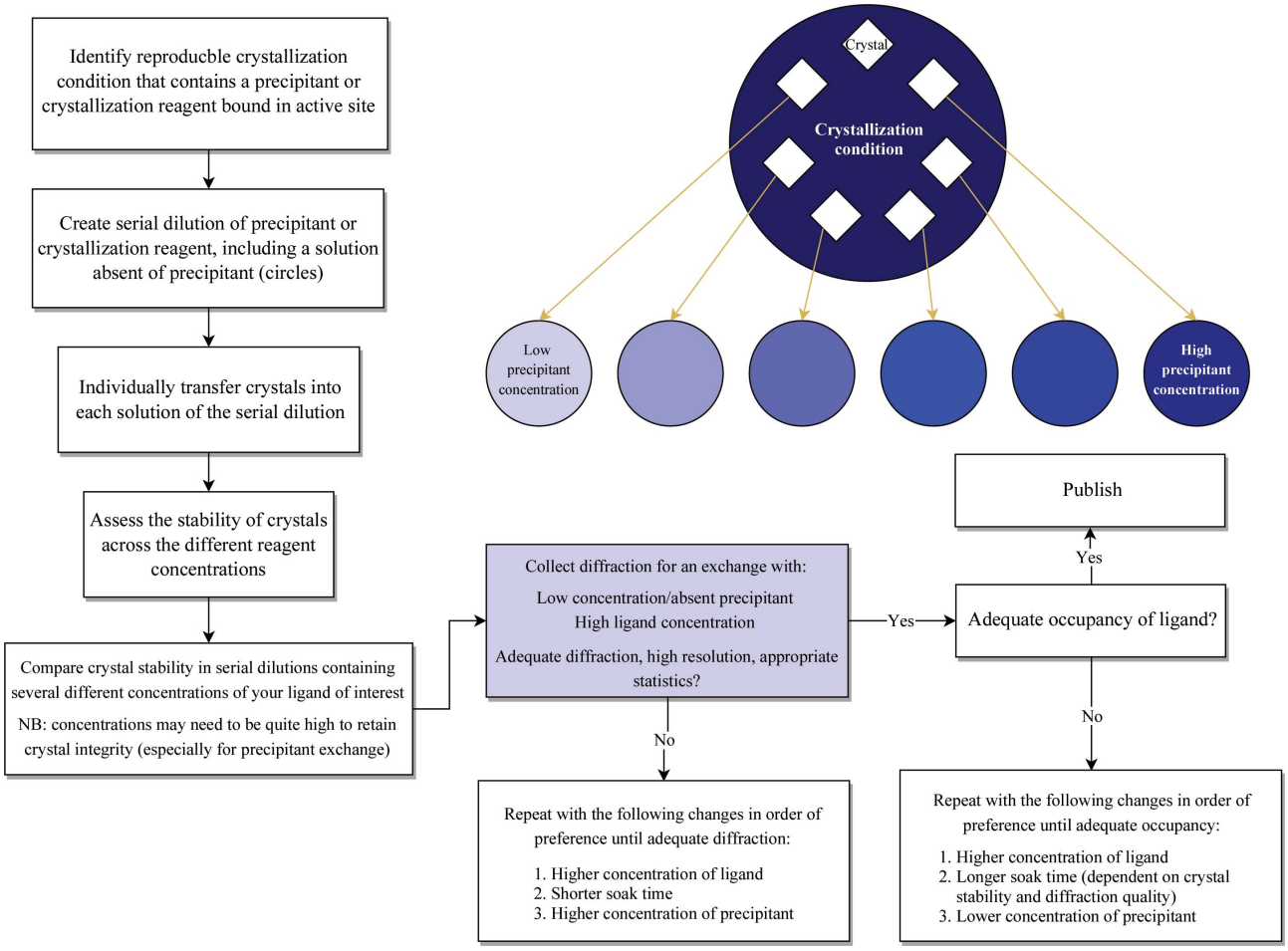



**Figure 6**
Flow diagram of the decision-making process during precipitant–ligand exchange.

structural investigation into the binding mode of ATP is ongoing. Interestingly, the $\beta$-phosphate of ADP in both conformations shared the same binding site as the $\beta$-phosphate of CTP within the P-loop. This was also the same site as observed for the sulfate precipitant, suggesting that this is a high-affinity binding site for anionic compounds. These structural data provide the first evidence for a molecule other than DAPA carbamate, or the closely related molecule 7-keto-8-aminopelargonic acid (KAPA), bound within the DAPA pocket. This revealed that the DAPA pocket can also accommodate larger and more varied compounds than the simple elongated acyl chains of DAPA and KAPA, a finding that will aid drug discovery. Finally, this structure indicates that inhibition by ADP may occur through competition with either substrate, a finding that invites further experimental investigation. The value of this structure lies in its contribution to mapping the binding characteristics of the active site of *Mt*DTBS. This will inform future antitubercular drug-discovery efforts targeting *Mt*DTBS activity and biotin biosynthesis, assisting in the design of inhibitors that bind with high specificity to the enzyme.

This technique may be of use in fragment-based drug discovery, where hits generally have weak affinity ($K_d$ > 500 µ*M*; Baker, 2013). If a fragment has been shown to bind to a protein using affinity and/or competitive inhibition assays, but only precipitant electron density is observed in the active site of the crystal structure, this technique can be used to exchange the precipitant, leading to fragment electron density. With weak ligands, it may not be possible to reduce the precipitant concentration to the required levels while maintaining crystal stability. This may lead to a scenario in which a certain proportion of protein molecules in the crystal are bound to precipitant, with the remainder bound to ligand. In this case, both precipitant- and ligand-bound states could be modelled (Pearce, Krojer & von Delft, 2017), or it may be useful to combine this technique with a program such as *PanDDA* (Pearce, Krojer, Bradley *et al.*, 2017) or polder maps (Liebschner *et al.*, 2017) to aid in discerning fragment density.

## 5. Conclusions

Here, we describe a crystallographic method for simultaneously removing and replacing active-site-bound precipitant molecules with a ligand of interest: a process that we have termed precipitant–ligand exchange. This technique was used to solve two CTP–*Mt*DTBS structures from independent

crystal-growth conditions. In validation of precipitant–ligand exchange, the results were consistent with the published literature. Subsequently, we solved a previously intractable structure of ADP bound to *Mt*DTBS. This structure revealed an alternate binding mode to that previously reported for CTP, providing new insight into the biology of the enzyme, as well as informing antitubercular inhibitor-design efforts. Given the prevalence of crystallographic reagents modelled in entries in the PDB, we believe that this method will be broadly applicable and will benefit structure-based interaction-analysis projects that utilize crystallography.

## Acknowledgements

This research was undertaken on the MX1 beamline at the Australian Synchrotron, which is part of ANSTO.

## Funding information

AT is the recipient of a University of Adelaide Postgraduate Scholarship. This work was supported by funding from the Channel 7 Children's Research Foundation (Project 181614).

## References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Baker, M. (2013). *Nature Rev. Drug Discov.* **12**, 5–7.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Ciccone, L., Vera, L., Tepshi, L., Rosalia, L., Rossello, A. & Stura, E. A. (2015). *Biotechnol. Rep.* **7**, 120–127.
Dai, R., Geders, T. W., Liu, F., Park, S. W., Schnappinger, D., Aldrich, C. C. & Finzel, B. C. (2015). *J. Med. Chem.* **58**, 5208–5217.
Dai, R., Wilson, D. J., Geders, T. W., Aldrich, C. C. & Finzel, B. C. (2014). *Chembiochem*, **15**, 575–586.
Dey, S., Lane, J. M., Lee, R. E., Rubin, E. J. & Sacchettini, J. C. (2010). *Biochemistry*, **49**, 6746–6760.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Huang, W., Jia, J., Gibson, K. J., Taylor, W. S., Rendina, A. R., Schneider, G. & Lindqvist, Y. (1995). *Biochemistry*, **34**, 10985–10995.
Käck, H., Gibson, K. J., Lindqvist, Y. & Schneider, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5495–5500.
Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
Karplus, P. A. & Diederichs, K. (2015). *Curr. Opin. Struct. Biol.* **34**, 60–68.
Kontopidis, G., Andrews, M. J. I., McInnes, C., Cowan, A., Powers, H., Innes, L., Plater, A., Griffiths, G., Paterson, D., Zheleva, D. I., Lane, D. P., Green, S., Walkinshaw, M. D. & Fischer, P. M. (2003). *Structure*, **11**, 1537–1546.
Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C. & Adams, P. D. (2017). *Acta Cryst.* D**73**, 148–157.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
Öster, L., Tapani, S., Xue, Y. & Käck, H. (2015). *Drug Discov. Today*, **20**, 1104–1111.
Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M. & von Delft, F. (2017). *Nature Commun.* **8**, 15123.
Pearce, N. M., Krojer, T. & von Delft, F. (2017). *Acta Cryst.* D**73**, 256–266.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Salaemae, W. (2015). PhD thesis, pp. 70–74. University of Adelaide, Australia.
Salaemae, W., Azhar, A., Booker, G. W. & Polyak, S. W. (2011). *Protein Cell*, **2**, 691–695.
Salaemae, W., Booker, G. W. & Polyak, S. W. (2016). *Microbiol Spectrum*, **4**, https://doi.org/10.1128/microbiolspec.VMBF-0008-2015.
Salaemae, W., Yap, M. Y., Wegener, K. L., Booker, G. W., Wilce, M. C. J. & Polyak, S. W. (2015). *Tuberculosis*, **95**, 259–266.
Thompson, A. P., Sternicki, L. M., Wegener, K. L., Lu, W., Zuo, L., Booker, G. W., Polyak, S. W. & Li, Y. (2016). *Jiangsu J. Prevent. Med.* **27**, 257–261.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Woong Park, S., Klotzsche, M., Wilson, D. J., Boshoff, H. I., Eoh, H., Manjunatha, U., Blumenthal, A., Rhee, K., Barry, C. E., Aldrich, C. C., Ehrt, S. & Schnappinger, D. (2011). *PLoS Pathog.* **7**, e1002264.

# Chapter 4:

## The identification and development of compounds that bind to *Mt*DTBS with high affinity

## 4.1 Introduction

This chapter focuses on the discovery and development of compounds that bind to *Mt*DTBS, with a future goal of producing an inhibitor of both the enzyme and the pathogen *M. tuberculosis*. This work continues a previously established screening campaign against the *Mt*DTBS active site, where *in silico* screening with 57 cytidine analogues (Sigma Aldrich) and 93,904 commercially available and rule of three compliant fragments (UCSF ZINC library) generated a number of predicted hits (Wanisa Salaemae and Prof. Grant Booker, University of Adelaide) (Salaemae, 2015). In this screen, *in silico* hits were ranked according to predicted binding energy and 26 hits with more favourable predicted energy than CTP were purchased for further testing (11 cytidine analogues and 15 ZINC fragments). Preliminary surface plasmon resonance (SPR) binding experiments with 4 cytidine analogues and 7 fragment hits indicated weak affinity binding, while docking results predicted the hits bound to the active site. This chapter describes the use of X-ray crystallography to confirm specific ligand binding and provide detailed binding pose information. This technique, in combination with further SPR, was then used to guide chemical development of hits towards higher affinity *Mt*DTBS binding. The chemical development of hits presented in this chapter utilized the principles of fragment-based lead discovery (FBLD).

As discussed in the introduction (Chapter 1), FBLD is a promising strategy for the development of novel antibiotics. After the identification of hits, there are several different structure-guided methods for compound optimisation in FBLD, including growing, merging or linking compounds together (Lamoree and Hubbard, 2018). Growing involves the addition of functional groups to a hit in order to create additional contacts to the protein target. Alternatively, two or more compounds with overlapping or adjacent binding poses can be merged or linked together in order to generate higher affinity interactions. Progress in fragment development is often measured using the ligand efficiency (LE) metric. LE is defined as the ratio of the free energy of binding to the heavy atom count (HAC) of the molecule ($\Delta G/HAC$), with 'good' values considered to be equal to or greater than 0.4 (Hopkins *et al*., 2014). The first successful cases of FBLD involved fragment linking (Shuker *et al*., 1996), and gave a practical example of how this method of compound development can generate super-additive affinity improvements (Jencks, 1981). However, the composition of the linker moiety is crucial to achieving high affinity, and must be designed to not only interact favourably with the protein, but also to retain the binding poses of the original fragments (Ichihara *et al*., 2011). Furthermore, linkers should have appropriate rigidity to minimise entropic penalties upon binding. This can be problematic in practice and consequently fragment growing is more commonly reported (Erlanson *et al*., 2016). In this chapter, FBLD principles were used to grow a promising hit based on previous structural investigation of the *Mt*DTBS active site. Additionally, ligands bound to distal regions of the active site were identified, encouraging the use of linking to conjugate these ligands.

Previous biophysical and high-resolution structural characterisation of *Mt*DTBS generated in Chapter 2 and 3, as well as previous literature, provided useful information for optimisation and development of fragment hits and linked compounds. Investigation into the ligand binding characteristics of the cytidine binding site, the phosphate binding loop (P-loop) and the DAPA pocket (Figure 4.1) revealed interactions that could be exploited during compound development to attain high affinity for *Mt*DTBS. In Chapter 3, the high affinity nucleoside triphosphate (NTP) binding mode was described. The cytosine moiety binds in a shallow binding pocket, making four hydrogen bonds with *Mt*DTBS residues. This highly efficient and specific interaction could be exploited during drug discovery. During the development of

the precipitant-ligand exchange technique presented in Chapter 2, it became apparent that the *Mt*DTBS P-loop was a high affinity binding site for anionic compounds. This was further investigated in Chapter 3, where it was found that the LRLR nest in particular exhibits extremely high affinity for phosphates, such as those present in nucleoside polyphosphates. This information suggested anionic interactions in this vicinity could also be exploited for drug discovery. Finally, structures of *Mt*DTBS in complex with DAPA-carbamate (PDB IDs: 3FMF, 6CVE) and ADP (Chapter 2) provided insights for the growing or merging of compounds bound to the largely hydrophobic DAPA pocket.

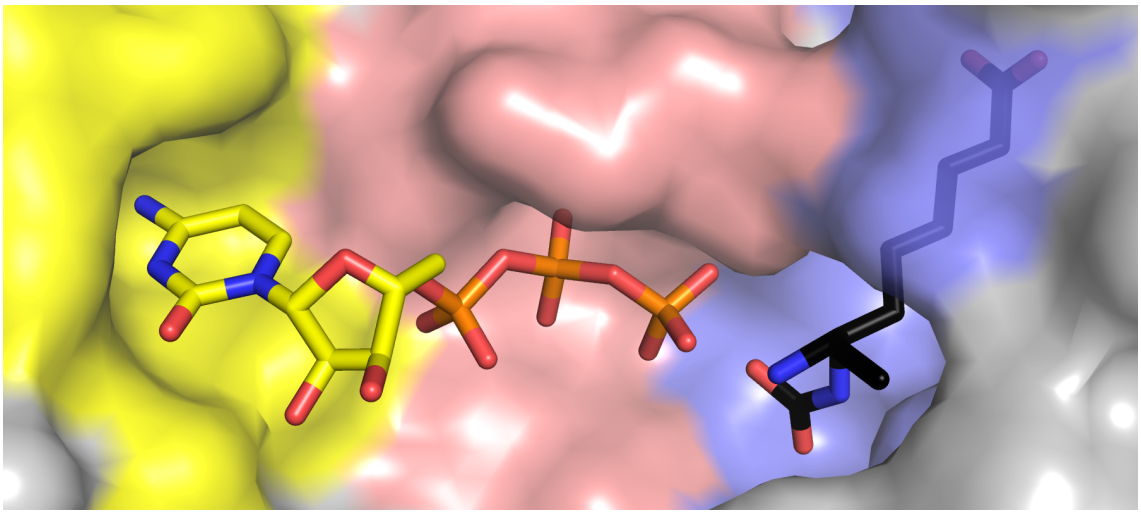


**Figure 4.1:** The active site of *Mt*DTBS in complex with CTP and DAPA-carbamate (PDB ID: 6CVE; Chapter 3). The surface of *Mt*DTBS (grey) is depicted with the following ligand-binding regions highlighted: the cytidine binding pocket (yellow), the P-loop (salmon) and the DAPA pocket (blue). CTP (yellow sticks) and DAPA-carbamate (black sticks) are displayed bound to the active site.

This chapter describes X-ray crystallographic binding confirmation of two cytidine analogues and one fragment that bound to the DAPA pocket, which were identified as promising starting points for compound growing and linking. The cytidine analogues were found to bind to the nucleoside binding pocket, and while the cytosine base appeared to be optimal, modifications at the 2′ ribose site were observed to bind, with one such modification resulting in improved affinity over cytidine. The structure solved from a fragment soak with *Mt*DTBS suggested that a degradation product of the original fragment was bound to the DAPA pocket of *Mt*DTBS. The chemical structure of this compound was confirmed through several orthogonal methods. Serendipitously, this compound constituted a convenient scaffold for structure guided fragment growing, enabling the development of a compound with greatly improved binding affinity for *Mt*DTBS. Finally, having identified ligands that bound to the cytidine binding site and the DAPA pocket, these compounds were linked and the binding affinity and mode analysed. Although these linked compounds did not improve binding over their substituents, X-ray crystallographic data revealed strategies that could be implemented to improve binding. This work constitutes the basis for future development of inhibitors of *Mt*DTBS.

## 4.2 Materials and Methods

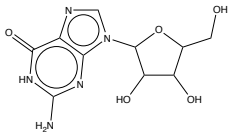
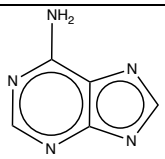
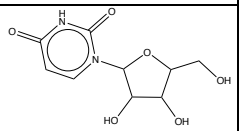
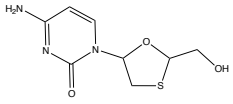
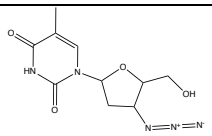
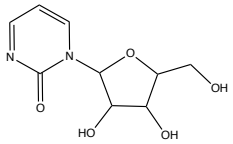
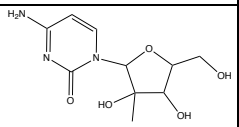
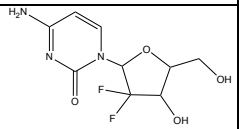
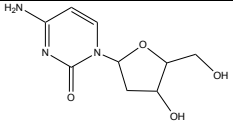
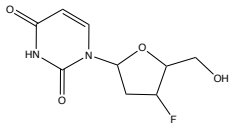
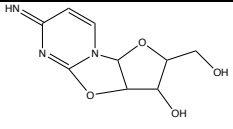
### 4.2.1 General X-ray crystallographic data collection and processing

Purified *Mt*DTBS protein was obtained as indicated in Chapters 2 and 3. Generally, ligands were soaked at as high concentration possible with consideration of the stock concentration. DMSO dissolved ligands were generally soaked such that a maximum concentration of 10% DMSO was allowed. Ligand treated *Mt*DTBS crystals were flash cooled in liquid nitrogen and subjected to X-ray diffraction on the MX-1 Beamline at the Australian Synchrotron, part of ANSTO (McPhillips *et al*., 2002). Data were collected at 100K and at a wavelength of 0.9537 Å. Diffraction data for **1**-CTN-*Mt*DTBS were collected at the Braggs Facility (University of Adelaide) using a Rigaku R-AXIS IV++ detector at a wavelength of 1.54 Å and at 100K. Datasets were indexed, scaled and merged in either iMosflm (Battye *et al*., 2011) or XDS (Kabsch, 2010) and AIMLESS (CCP4) (Winn *et al*., 2011) in accordance with $CC_{1/2}$ cut-off values (Karplus and Diederichs, 2012, 2015). The phase problem was solved via molecular replacement (PhaserMR (McCoy *et al*., 2007) or DIMPLE) using in-house *Mt*DTBS search models. Structures were modelled through iterative rounds of building in Coot (Emsley and Cowtan, 2004) and refinement in Phenix.refine (Adams *et al*., 2010).

### 4.2.2 Crystallographic screening of *in silico* hits

Select *in silico* hits were subjected to crystallographic screening against *Mt*DTBS (ligand purchasing and formulation details in (Salaemae, 2015)). Cytidine analogue hits and alternate nucleosides (Table 4.1) were screened via crystal soaking using the methodology provided in 4.2.1, however, in the screening pipeline, each dataset for the cytidine analogue screen was solved using AIMLESS and DIMPLE to the maximum resolution which resulted in $CC_{1/2}$ and completeness values above 0.9 and 90% respectively. Each model was then subjected to PanDDAs refinement (Pearce *et al*., 2017) and hits were manually inspected for ligand electron density.

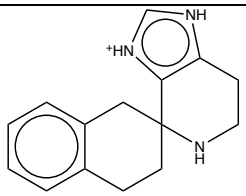
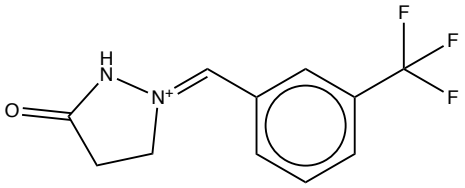
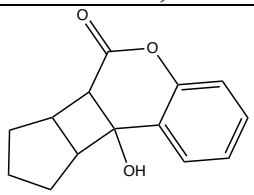
**Table 4.1:** Lab denotation, name, structure and concentration used in the crystallographic screening of cytidine analogue hits and select nucleosides. The bolded compounds indicate that the crystal structure contained electron density pertaining to the ligand described.

| Lab book denotation | Compound name and structure | Concentration (mM) | Lab book denotation | Compound name and structure | Concentration (mM) |
|---|---|---|---|---|---|
| - |  Guanosine | 100 | - |  Adenine | 100 |
| - |  Uridine | 100 | CT1 |  Lamivudine | 21 |
| CT2 |  3'-Azido-3'-deoxythymidine | 17 | CT4 |  Zubularine | 6 |
| CT5 |  2'-C-Methylcytidine | 53 | **CT6* (dFdC)** |  **Gemcitabine hydrochloride** | **56** |
| **CT7* (dCTN)** |  **2'-Deoxycytidine hydrochloride** | **~50 (slurry)** | CT10 |  2',3'-dideoxy-3'-fluorouridine | 15 |
| CT12 |  Ancitabine hydrochloride | 167 | | | |

*Crystallography described in more detail in table 4.4

Fragment hits (Table 4.2) were tested via a combination of co-crystallography (incubating the ligand with the protein prior to crystal growth via vapour diffusion), and traditional soaking (adding ligand directly to drops containing pre-grown crystals). Several fragment hit datasets were solved in accordance with the methodology provided in 4.2.1 before examination of the structure for electron density pertaining to the ligand. More detailed soaking methodology involving compound **1** is provided in Table 4.4. Compound **1** was purchased from Specs.

**Table 4.2:** Lab book denotation, ZINC ID, name, structure and details of the attempts to generate a complex by crystallography for three of the previously determined *in silico* hits. The bolded crystallographic attempt indicates that the crystal structure contained ligand electron density.

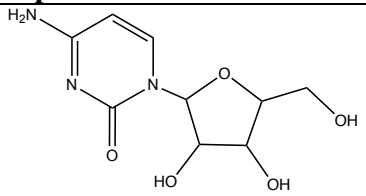
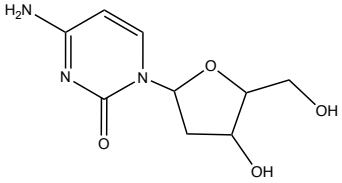
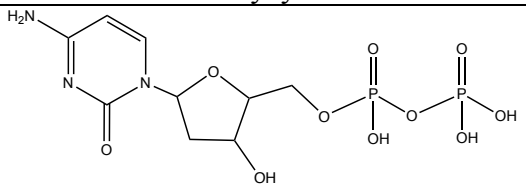
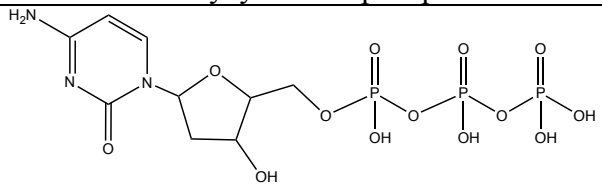
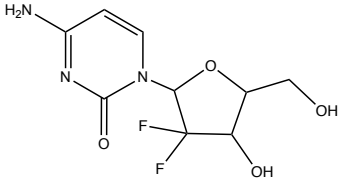
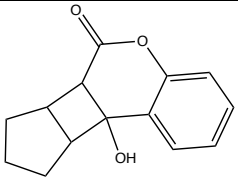
| Lab book denotation | Zinc ID | Compound name and structure | Crystallography attempts |
|---|---|---|---|
| B1 | 65409233 |  (3aR,4S)-spiro[3a,5,6,7-tetrahydroimidazo[4,5-c]pyridine-4,2'-tetralin] | Co-crystallized at 10 mM<br><br>Soaked at 45 mM |
| B7 | 12956616 |  1-(3-Trifluoromethylphenyl)methyleneimmonium-3-pyrazolidine-1,2-inner salt | Co-crystallized at ~2 mM<br><br>Soaked at ~5 mM |
| B9<br>(**1**) | 04114260 |  9b-hydroxy-6b,7,8,9,9a,9b-hexahydrocyclopenta[3,4]cyclobuta[1,2-c]chromen-6(6aH)-one | Co-crystallized at 10 mM<br><br>**Soaked at 30 mM*** |

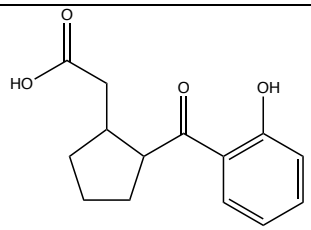
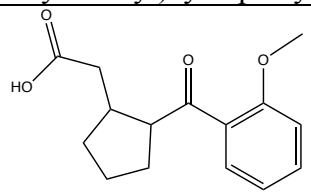
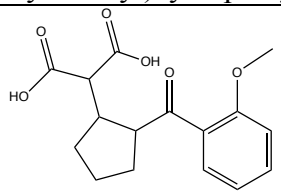
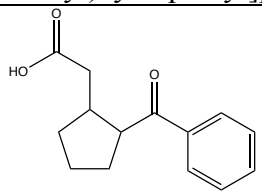
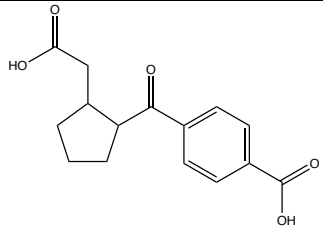
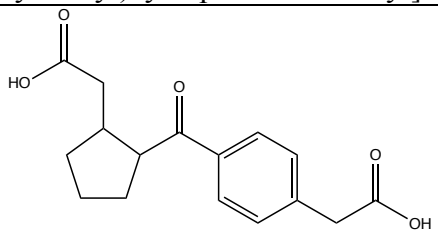
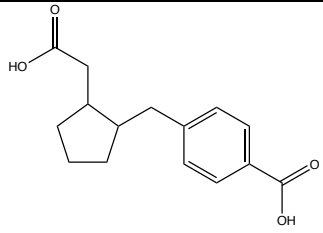
*Crystallography described in more detail in table 4.4

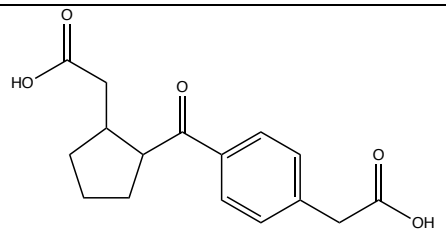
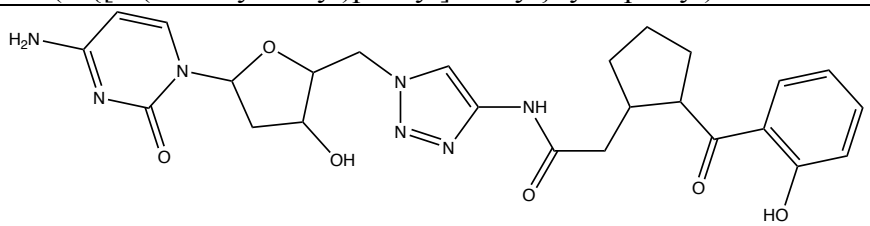
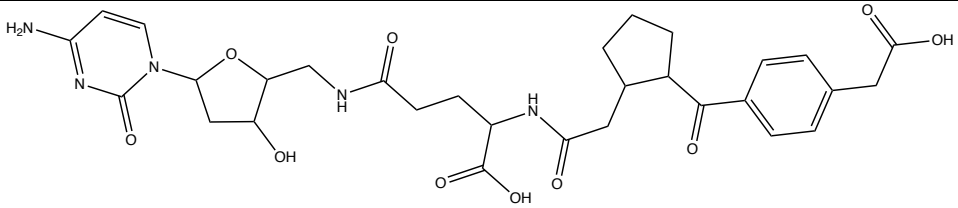
### 4.2.3 Confirmed hits and their analogues

Following X-ray crystallographic determination of the binding mode of cytidine analogues and fragment hits, analogues were identified and synthesized for structure activity relationship (SAR) studies. Table 4.3 provides a list of all chemical structures referred to in this chapter. Details of their use in *Mt*DTBS crystallography studies are provided in Table 4.4. Cytidine, its analogues, and compound **1** are listed again for ease of structural comparison.

**Table 4.3:** Abbreviations, names and structures of crystallographic hits, analogues of hits, synthesized compounds and other compounds referred to in Chapter 4.

| Abbreviation | Compound name and structure |
|---|---|
| CTN |  cytidine |
| dCTN |  2′-deoxycytidine |
| dCDP |  2′-deoxycytidine diphosphate |
| dCTP |  2′-deoxycytidine triphosphate |
| dFdC |  2′,2′-difluoro-2′-dexoycytidine (a.k.a. Gemcitabine) |
| **1** |  9b-hydroxy-6b,7,8,9,9a,9b-hexahydrocyclopenta[3,4]cyclobuta[1,2-c]chromen-6(6aH)-one |

| | |
|---|---|
| **2** |  <br>2-(2-(2-hydroxybenzoyl)cyclopentyl)acetic acid |
| **3** |  <br>2-[2-(2-methoxybenzoyl)cyclopentyl]acetic acid |
| **3a** |  <br>2-[2-(2-methoxybenzoyl)cyclopentyl]propanedioic acid |
| **4** |  <br>2-(2-benzoylcyclopentyl)acetic acid |
| **5** |  <br>4-[2-(carboxymethyl)cyclopentanecarbonyl]benzoic acid |
| **6** |  <br>2-{2-[4-(carboxymethyl)benzoyl]cyclopentyl}acetic acid |
| **7** |  <br>4-{[2-(carboxymethyl)cyclopentyl]methyl}benzoic acid |

| | |
|---|---|
| **8** |  <br>2-(2-{[4-(carboxymethyl)phenyl]methyl}cyclopentyl)acetic acid |
| **9** |  <br>N-[(1-{[(2R,3S,5R)-5-(4-amino-2-oxo-1,2-dihydropyrimidin-1-yl)-3-hydroxyoxolan-2-yl]methyl}-1H-1,2,3-triazol-4-yl)methyl]-2-[(1S,2R)-2-(2-hydroxybenzoyl)cyclopentyl]acetamide |
| **10** |  <br>4-({[5-(4-amino-2-oxo-1,2-dihydropyrimidin-1-yl)-3-hydroxyoxolan-2-yl]methyl}carbamoyl)-2-(2-{2-[4-(carboxymethyl)benzoyl]cyclopentyl}acetamido)butanoic acid |

4.2.4 Preparation of compounds found in Table 4.3

Gemcitabine, 2′-deoxycytidine, 2′-deoxycytidine diphosphate and cytidine were purchased as analytical grade powders from Sigma Aldrich and solubilised in water. The stock concentration was determined according to Beer's Law, using literature wavelength and extinction coefficient values (Sambrook *et al.*, 1989) and UV absorption (Nanodrop 2000, Thermo Scientific).

Compounds **2** (subsequent to initial discovery) and **4** were synthesized by Birgit Gaiser (Department of Chemistry, University of Adelaide). Compounds **3**, **5**-**9** were synthesized by Kwang Jun Lee (Department of Chemistry, University of Adelaide). Concentrated stock solutions of synthesized compounds were prepared in 100% DMSO (Sigma Aldrich) (100-330 mM).

4.2.5 Crystallography methodology for new structures

All *Mt*DTBS crystals were grown via the hanging drop method (1:1 drop:well ratio, 2-4 uL total volume) with a 500 µL well solution in either of the following conditions: (1) 1.2 – 1.7 M **ammonium sulfate**, 0.1 M Tris pH 8, 10-15% glycerol (2) 0.7 M **sodium citrate**, 0.1 M Tris pH 7 and 20% glycerol. Co-crystallisation involved growing crystals in the presence of the ligand. Soaking was performed by adding the desired ligand directly to the crystallization condition containing a pre-grown crystal. Precipitant ligand exchange involved transferring crystals into a solution where ligand replaced the precipitant (highlighted above in bold). Crystallographic data tables are provided in Appendix 1.

**Table 4.4:** Crystallisation and ligand complexing methodology for all new structures.

| Ligand | Condition | Concentration | Other | Method | Time | Cryoprotection |
|---|---|---|---|---|---|---|
| dCTN | 1 | 56 mM | - | Soaking | 8 days | - |
| dCDP | 2 | 20 mM | 2 mM $MgCl_2$ | Co-crystallisation | - | Paratone-N |
| dCTP | 2 | 10 mM | 1 mM $MgCl_2$ | Soaking | Overnight | |
| dFdC | 1 | ~50 mM (slurry) | - | Soaking | 2 days | - |
| **2 (1*)** | 1 | 30 mM | 10% DMSO | Soaking | 8 days | - |
| **Synthesized-2** | 1 | 20 mM | 10% DMSO | Soaking | 2 days | - |
| **2** + CTN | 1 | 30 mM (1)<br>60 mM (CTN) | 5% DMSO | Soaking | Overnight | - |
| **3a** | 2 | 10 mM | 10% DMSO | Precipitant ligand exchange | 2 days | - |
| **4** | 1 | 30 mM | 10% DMSO | Soaking | 2 days | - |
| **5** | 1 | 10 mM | 10 % DMSO | Soaking | Overnight | - |
| **6** | 1 | 10 mM | 10 % DMSO | Soaking | Overnight | - |
| **7** | 1 | 10 mM | 10 % DMSO | Soaking | Overnight | - |
| **8** | 1 | 10 mM | 10 % DMSO | Soaking | Overnight | - |
| **9** | 1 | 10 mM | 10% DMSO | Precipitant ligand exchange | < 30 min | - |
| **10** | 2 | 10 mM | 10% DMSO | Soaking | 2 days | - |

*NB: **2**-*Mt*DTBS crystal structure was obtained by soaking with compound **1**.

### 4.2.6 Surface plasmon resonance binding experiments

SPR binding analysis was performed using a Biacore S200 instrument (GE Healthcare). Biotinylated-*Mt*DTBS was immobilized via streptavidin mediated coupling as described in Chapter 3. Streptavidin alone was used in the reference cell, for reference subtraction in all analyses. Immediately before analysis, compounds were diluted into 50 mM Tris pH 7.4, 150 mM NaCl, 0.05% Tween20 to a final concentration of 5% DMSO. Compounds were then further diluted to the maximum test concentrations and serially diluted in running buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.05% Tween20, 5% DMSO). 8-point solvent correction curves were obtained using 50 mM Tris pH 7.4, 150 mM NaCl, 0.05% Tween20 with DMSO ranges of 4.2% – 5.7%. Solvent correction curves were used as the secondary correction for all binding analyses for compounds that were dissolved in DMSO. SPR results and example sensorgrams for binding measurements with compounds **2**-**10**, dCDP and dCTP are provided in Appendix 2.

## 4.3 Results

### 4.3.1 General description of *Mt*DTBS crystal structures

Crystallographic methodology for all new structures discussed in this chapter is provided in Materials and Methods 4.2.1, 4.2.2 and 4.2.5 (Table 4.4) and crystallographic data tables for all new datasets are provided in Appendix 1. All *Mt*DTBS crystal structures described in this chapter were solved in the $P2_12_12_1$ space group, with identical overall fold to previous structures (Dey et al., 2010; Salaemae et al., 2015; Thompson et al., 2018b, 2018a). The asymmetric unit contained a dimer of the catalytic *Mt*DTBS homodimer, which was comprised of coil regions that connected alternating α-helices and β-sheets (Figure 4.2). Notable structural features of the active site include the cytidine binding site (formed by the coils connecting β7-$3_{10}2$ and β6-α7), the phosphate binding loop (P-loop; between α1 and β1) and the DAPA pocket (a pocket formed at the homo-dimerisation interface) (compare to Figure 4.1 for surface representation).

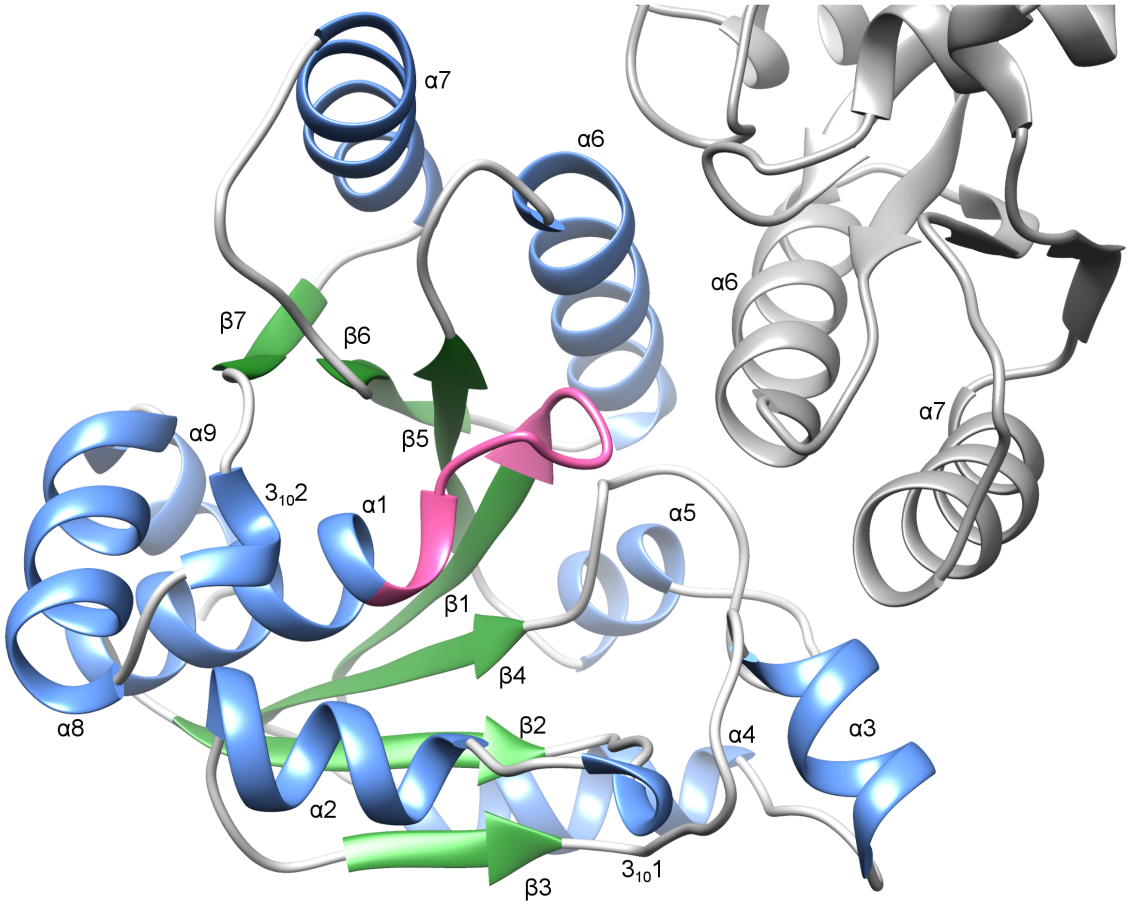



**Figure 4.2:** Representative overall fold of the catalytic *Mt*DTBS homodimer in crystal structures. Each monomer contains α-helices (blue) and β-sheets (green). Dimerization occurs along an interfaced formed by α3, α6 and α7 of the neighbouring subunit (grey). The phosphate binding loop (pink) is highlighted to visualize the active site.

### 4.3.2 Investigation into *in silico* cytidine analogue hits

In order to investigate the hits from the *in silico* screening described in section 4.1, the 8 available cytidine analogues were subjected to crystallographic soaking experiments with *Mt*DTBS (Table 4.1). To further probe the active site, guanosine, adenine and uridine were also included. From this experiment, 11 datasets were solved according to the crystallographic methodology detailed in 4.2.1; 4.2.2; 4.2.5. Of these 11 datasets, ligand electron density was only observed in the cytidine binding site for *Mt*DTBS crystals soaked with 2′-deoxycytidine (dCTN) and 2′,2′-difluoro-2′-dexoycytidine (aka gemcitabine, denoted dFdC). These molecules differ from cytidine by the substituents present at the 2′ position on the ribose ring, where cytidine contains a 2′-OH, dCTN contains two protons, and dFdC contains two fluorine atoms (Table 4.1, 4.3; Figure 4.3B, C). The hydrogen bonding arrangement between *Mt*DTBS and the cytosine bases were identical to that previously described in Chapter 3 for CTN. These interactions involved cytosine $4NH_2$ and the backbone carbonyl of Pro197 and Gly169, cytosine 3N and the backbone amide of Ala200, and hydrophobic interactions between cytosine 4C and 5C, and the side chain of Val17 (Figure 4.3). Compounds that varied at the cytosine moiety were not observed by crystallography, consistent with the notion that the cytosine interactions are highly optimized for binding to this pocket. The absence of alternate nucleoside interactions is also consistent with the promiscuous ligand binding model (Chapter 3). Adenine was not observed in the DAPA pocket, as found in the ADP-*Mt*DTBS structure in Chapter 2. This was likely because the affinity of the adenine moiety is substantially lower without the ribose and diphosphate.

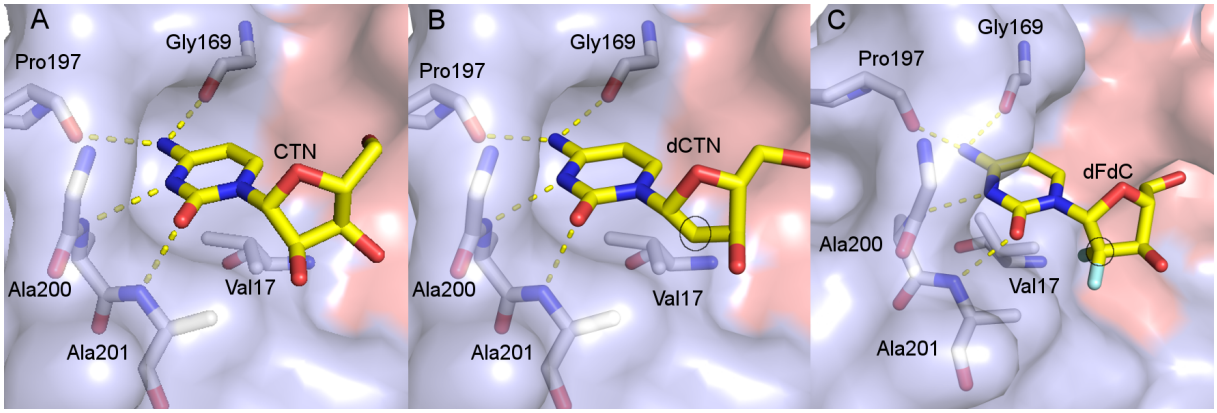



**Figure 4.3:** The *Mt*DTBS active site (blue surface; interacting residues represented by grey sticks) bound to cytidine and analogues (yellow sticks). **A.** cytidine (CTN) **B.** 2′-deoxycytidine (dCTN), **C.** gemcitabine (2′,2′-difluoro-2′-dexoycytidine, dFdC). Both cytidine analogues adopt similar poses, and interact with the residues Gly169, Pro197, Ala200 and Ala201 as previously demonstrated for CTN. The substituents present at the 2′ position on the ribose ring of the analogues (circled) are proximal to the hydrophobic sidechains of residues Ala201 and Val17.

The cytidine binding site was able to accommodate chemical variation at the 2′ ribose position of cytidine. This site is adjacent to the hydrophobic residues Ala201 and Val17. Therefore, it was hypothesized that the 2′ protons of dCTN or fluorines of dFdC may interact more favourably with *Mt*DTBS than the more hydrophilic 2′OH of cytidine (Figure 4.3). SPR binding experiments were unable to distinguish the affinities for dCTN and dFdC, due to their weak binding ($K_D$ > 1 mM). However, 2′-deoxycytidine diphosphate (dCDP; $K_D$ = 120 ± 30 nM, n = 3) bound approximately 3-fold tighter than CDP ($K_D$ = 450 nM ± 50, n = 4), and 2′-deoxycytidine triphosphate (dCTP; $K_D$ = 47 nM, n = 3) likewise bound with ~3-fold higher affinity than CTP ($K_D$ = 160 ± 10 nM, n = 5) (Figure 4.4A, B; SPR data tables presented in Appendix 2). X-ray crystallographic investigation of the dCDP and dCTP binding modes revealed an identical binding pose to that observed for CDP and CTP (Figure 4.4C), demonstrating these molecular composition changes don't alter binding mode. SPR data were fit using kinetic binding analysis (Biacore S200) to determine dCDP association and dissociation rates ($k_a = 1.2 \pm 0.4\times10^6$ $M^{-1}s^{-1}$; $k_d = 1.2 \pm 0.3\times10^{-1}$ $s^{-1}$). Comparison to CTP ($k_a = 3.2 \pm 0.3\times10^5$ $M^{-1}s^{-1}$, $k_d = 5.3 \pm 0.4\times10^{-2}$ $s^{-1}$) revealed that dCDP associated approximately 4-fold faster than CTP, but also dissociated 2-fold faster, explaining the similarity in $K_D$ values. While dCTP exhibited association rates between those of dCDP and CTP, it dissociated the slowest out of all cytidine or deoxycytidine polyphosphates ($k_a = 7.13 \pm 0.05\times10^5$ $M^{-1}s^{-1}$; $k_d = 3.4 \pm 0.02\times10^{-2}$ $s^{-1}$). A graphical representation of binding kinetics values for CTP, CDP, dCDP and dCTP is provided in Figure 4.4B. These data demonstrate that the 2′ hydroxy group leads to a slower association rate and faster dissociation rate, culminating in the observation that dCTP is the preferred NTP when measured by binding affinity. For the purposes of linked inhibitor design, dCTN currently constitutes the highest affinity ligand for the cytidine binding site. Further investigation is required to determine the effect that the 2′-difluoro group has on the binding of dFdC, as well as to explore other chemical variations at this position.

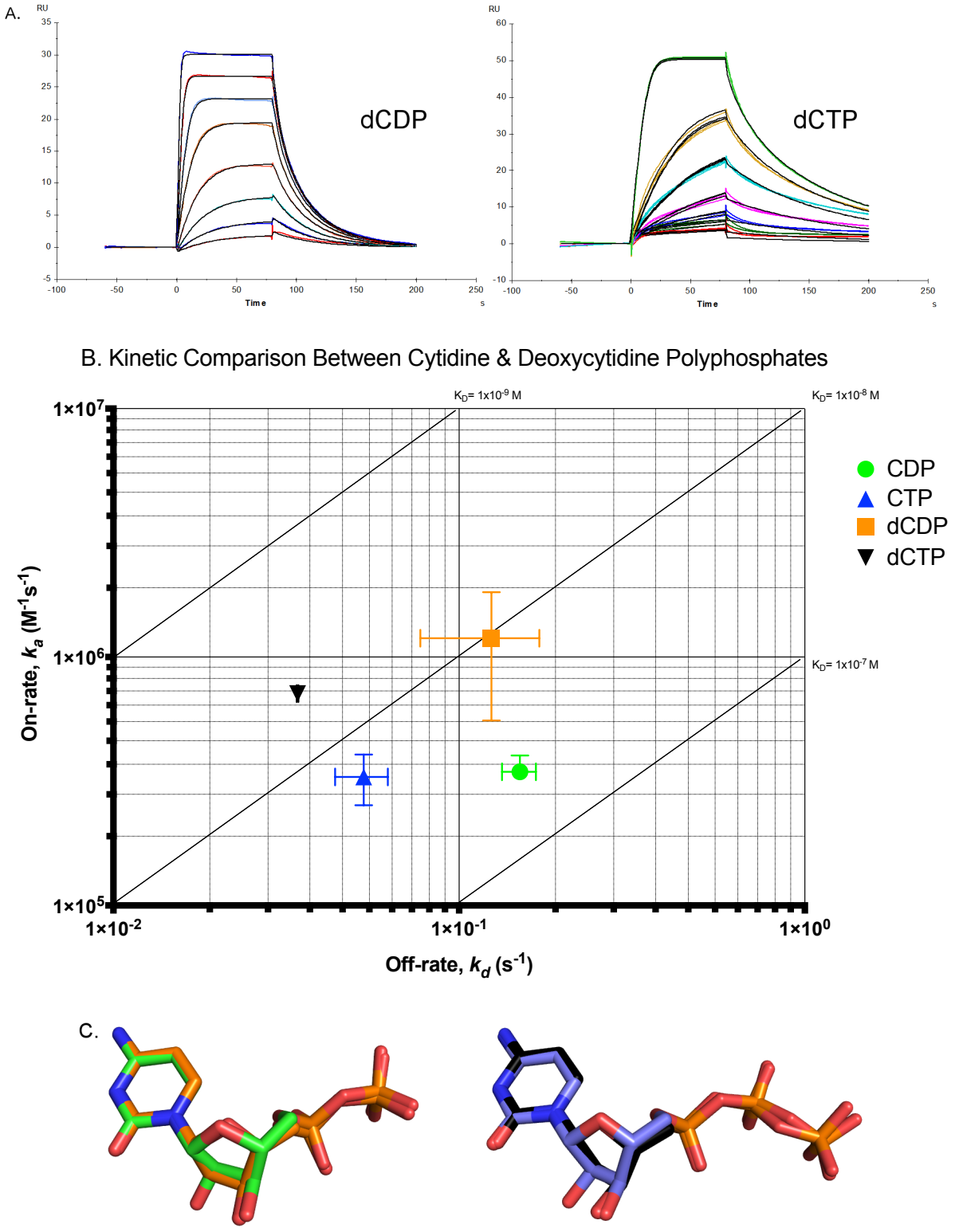



**Figure 4.4: A.** Examples of SPR sensorgrams for dCDP and dCTP binding to *Mt*DTBS. Both ligands exhibited slow association and dissociation rates and were accordingly fit with the kinetic binding model (Biacore S200 Evaluation Software). **B.** Comparison of association and dissociation rates for CDP (green circle), CTP (blue triangle), dCDP (orange square) and dCTP (black triangle). Error bars represent standard error of the mean. dCDP associated and dissociated faster than CTP, resulting in similar $K_D$ values (diagonal lines). dCTP had the slowest dissociation of all polyphosphates, contributing to its high affinity. NB: error bars were smaller than the size of the symbol for dCTP **C.** Overlays of the binding poses of CDP (green sticks) and dCDP (orange sticks), as well as CTP (blue sticks) and dCTP (black sticks). Each molecular pair adopted essentially identical binding poses.

### 4.3.3 Investigation into *in silico* fragment hits

**X-ray crystallography reveals an unexpected molecule bound to the DAPA pocket**
Three fragment hits were available for binding mode investigation via X-ray crystallography; data were collected for several soaking experiments with *Mt*DTBS crystals (Methodology in 4.2.1 – 4.2.5). Electron density corresponding to a ligand bound to the active site was only observed for one structure, which will be the sole focus of this section. Ligand electron density was observed in the DAPA pocket, for a structure solved from crystals soaked with 9b-hydroxy-6b,7,8,9,9a,9b-hexahydrocyclopenta[3,4]cyclobuta[1,2-c]chromen-6(6aH)-one (**1**; 30 mM final concentration; X-ray crystallographic data in Appendix 1). Interestingly, this electron density was not consistent with that expected for compound **1** (Figure 4.5A), and low density fit scores (Coot; 0.5-0.7) and real-space correlation coefficients (PHENIX; 0.565-0.814) were determined for this ligand after modelling and refinement. The observed density was also inconsistent with other crystallographic reagents present, i.e. sulfate, glycerol or DMSO, implying it may instead correspond to a degradation product of **1**.

The presence of a phenolic lactone, fused to a strained cyclobutane ring, suggested that compound **1** may be susceptible to hydrolysis and subsequent ring opening to form 2-(2-(2-hydroxybenzoyl)cyclopentyl)acetic acid (**2;** Figure 4.5B). Compound **2** clearly exhibited higher visual coherence with the observed electron density than molecule **1** (Figure 4.5C, D), as well as high density fit scores (Coot; > 1.09) and real-space correlation coefficients (PHENIX; > 0.938). Predicted compound **2** contained two stereocenters, of which the stereochemistry was unknown. While compound **1** was provided as a racemic mixture of the *cis*,*cis*,*cis*-isomer (Specs, Netherlands), racemisation of **2** was likely to occur during hydrolysis due to the relatively acidic nature of the protons adjacent to the ketone. Although trans stereochemistry at the cyclopentane ring was hypothesized to be more thermodynamically stable, each of the four possible diastereomers (RS/SR/RR/SS) were simultaneously modelled into the crystal structure to see which most closely matched the observed electron density (Figure 4.5C, D). Trans diastereomer RS-**2** exhibited the highest occupancy after refinement (RS: 0.51, SR: 0.13, RR: 0.32, SS: 0.34) and the most preferable real-space refinement statistics. However, due to the minimal structural differences observed between diastereomers (Figure 4.5C, D), it was not possible to distinguish the stereochemistry of **2** with confidence. Thus, the crystallographic modelling was consistent with a stereoisomer of **2** bound to *Mt*DTBS, and most likely the RS form, although confirmatory experimental evidence was required.

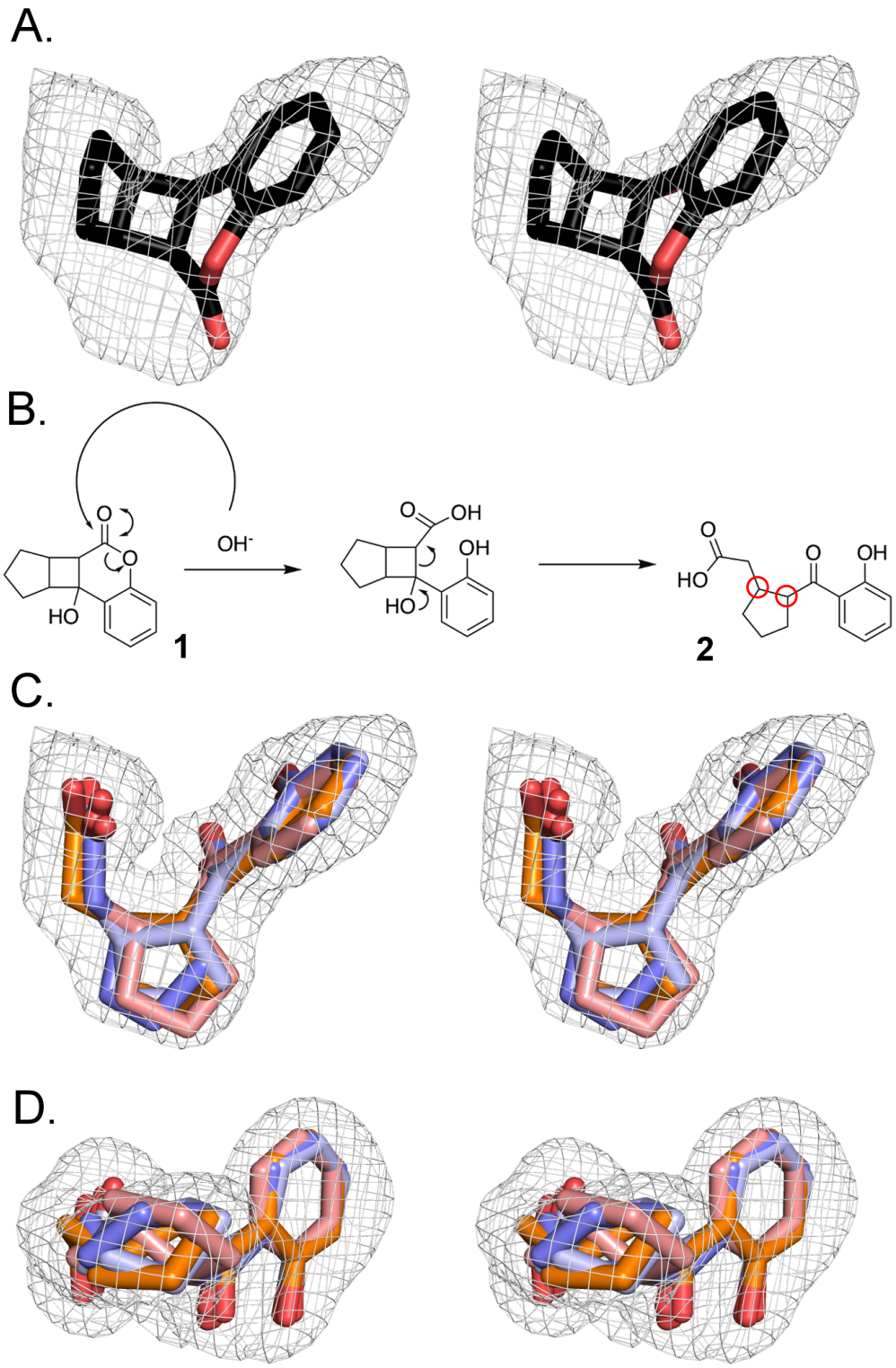



**Figure 4.5: A.** Walleye view of **1** (black sticks) within Polder electron density (3σ) observed within the active site of *Mt*DTBS. This molecule fits poorly into the density, and does not appear to be compatible with its structure. **B.** Proposed mechanism for the degradation of lactone **1** to carboxylic acid **2**. Degradation was hypothesized to occur via hydrolysis of the lactone and collapse of the four-membered ring. This degradation was accompanied by the generation of two new stereocentres (red circles). **C, D.** Two walleye perspectives of each diastereomer of **2** (RR: light blue, SS: dark blue, SR: pink, RS: orange) fit into Polder electron density (3σ) observed in the *Mt*DTBS active site. Each of these molecules is visually consistent with the electron density. Minimal structural differences between diastereomers were observed, and the preferred molecule could not be definitively identified solely by crystallography.

**Experimental confirmation of the molecular structure of 2**

The DMSO stock solution of compound **1** was analysed for the presence of degradation product **2** by a combination of small molecule X-ray crystallography, NMR and MS analysis (performed by Birgit Gaiser, University of Adelaide, Appendix 3) and LC/MS (Adelaide Proteomics Centre, Appendix 4). The vial with the stock of **1** contained a crystal, as well as an amorphous solid. This crystal was analysed by small molecule X-ray crystallography and was shown to be *cis*-*cis*-*cis*-**1** (Appendix 3). NMR analysis of the amorphous residue revealed that this sample consisted largely of **1**, but that resonances consistent with **2** were also present (Appendix 1). The ratio of compounds was estimated to be 3:1 in favour of **1** after accounting for the removal of the crystal prior to NMR analysis. Resonances from **2** indicated the presence of two different stereoisomers, and NOESY experiments revealed the major component was the *trans*-isomer, with only small amounts of the *cis*-isomer present (~75% *trans*-**2**; Appendix 3). Finally, LC/MS analysis of the stock of **1** indicated main ions, with *m/z* 231 ($C_{14}H_{15}O_3$) corresponding to **1**, 249 ($C_{14}H_{17}O_4$) corresponding to **2**. An additional ion with *m/z* 213 ($C_{14}H_{13}O_2$) was observed, consistent with an unknown compound similar to **1** with a loss of $H_2O$ (Appendix 4). This compound was also observed in MS of a freshly reconstituted stock of **1**, indicating that it is a by-product from compound synthesis. Together these results suggested that moisture present in the sample had induced partial degradation of **1** to **2**.

An independent sample of compound **2** was then intentionally produced by treatment of freshly purchased **1** with aq. LiOH (Birgit Gaiser, University of Adelaide). The NMR resonances of this compound were consistent with **2** and matched peaks in the spectrum of the original DMSO stock of **1** (Appendix 3)**.** MS similarly confirmed the presence of **2** (Appendix 3). Compound **2** was found to be in the *trans*-conformation only, leading to the conclusion that the *cis*-isomer of **2** is indeed unstable and conversion to the *trans*-isomer occurs rapidly. Attempts to separate the trans isomers through the use of a chiral auxiliary proved unsuccessful, and separation of the *trans*-**2** isomers for affinity comparison remains a future goal of this project. X-ray crystallographic data from *Mt*DTBS crystals soaked with hydrolysis product **2** revealed essentially identical active site electron density to that in the original crystal structure. Together, these data confirmed the identity of the original ligand as **2**, and that it was present in the original stock solution, most likely as the result of hydrolysis by trace amounts of water.

**The *Mt*DTBS crystallisation condition causes degradation from 1 to 2**

Soaking with **1** was originally carried out for 8 days in crystallisation condition 1, which is basic (pH 8, Section 4.2.3 Crystallography). As the solution used to deliberately hydrolyse **1** into **2** was also basic, it was hypothesized that further degradation to **2** could result from the crystallisation condition itself. A small amount of **1** was added to this solution and monitored by TLC over the course of a week. The first traces of **2** were detected after 2 days and complete conversion occurred after 6 days (Appendix 3). Therefore, even though some **2** may have been present in the stock solution, complete degradation of **1** likely occurred during the long soaking experiment in the crystallography condition.

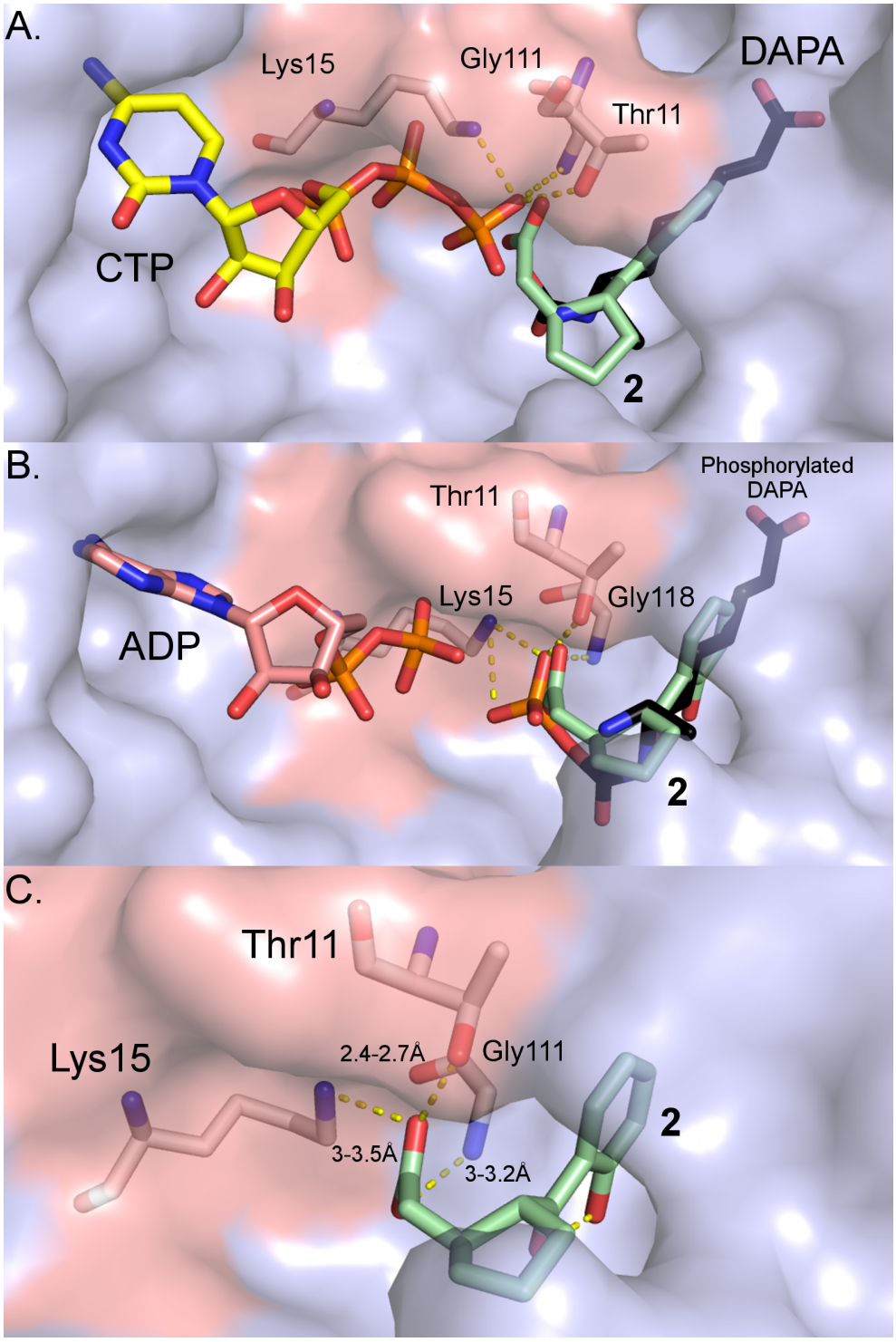



**Figure 4.6:** *Mt*DTBS (blue/pink surface) binding to **2** (green sticks) exhibits similar interactions to natural substrates. **A.** Overlay with CTP (yellow sticks) and DAPA (black sticks) (PDB ID: 6CVE). The carboxyl group of **2** occupies a similar region to the γ-phosphate of CTP bound at the P-loop (pink surface). Specific polar contacts to Thr11, Lys15 and Gly111 are depicted. The remainder of **2** binds in the DAPA pocket, however DAPA extends further into this binding pocket with its carboxyl group. **B.** Overlay with ADP (pink sticks) and the mixed carbamic phosphoric acid anhydride of DAPA (black sticks) (*Ec*DTBS; PDB ID: 1DAK). The carboxyl group of **2** overlays well with the phosphoryl group, which also exhibits polar interactions with Thr11, Lys15 and Gly118 (the *Ec*DTBS equivalent residue of Gly111). This indicates that inhibitors developed from **2** could function as mimics of the reaction intermediate. **C.** Binding of **2** is favoured by hydrogen bonds from the carboxyl group of **2** to the side chains of the P-loop residues Thr11 and Lys15, as well as an additional interaction to the amide backbone of Gly111.

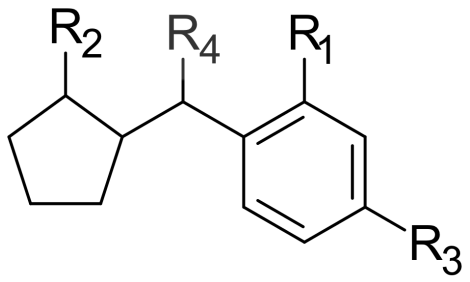
**2 mimics the phosphorylated reaction intermediate**

Compound **2** was observed bound to the *Mt*DTBS DAPA pocket. Superposition with the structure of *Mt*DTBS in complex with CTP and DAPA-carbamate (PDB ID: 6CVE) revealed that **2** replicates interactions exhibited by the biological substrates and the enzyme (Figure 4.6A, B). The **2** carboxy group occupied a similar region to the γ-phosphate of CTP, replicating hydrogen bonding interactions to the side chain of Thr11 (2.36 – 2.71Å), Lys15 (2.96 – 3.45Å), amide backbone of Gly111 (3.02 – 3.21Å), and several water molecules (side chain interactions depicted in Figure 4.6C). However, comparison to the structure of *E. coli* DTBS bound to a trapped reaction intermediate, revealed that the **2** carboxyl group more closely mimics the phosphate of the mixed carbamic phosphoric acid anhydride of DAPA, formed by phosphoryl transfer from ATP to DAPA-carbamate (Figure 4.6B; PDB ID: 1DAK). The remainder of **2** binds to the hydrophobic DAPA pocket, with the pentane and aryl rings occupying the same region as the DAPA 8-amino group and alkyl chain respectively (Figure 4.6A, B). SPR measurements with **2** revealed it binds with weak affinity ($K_D$ = 3.4 ± 0.4 mM; Table 1; Appendix 2). This was also reflected by the requirement for high concentrations for the crystal soaking, as electron density for **2** was not observed in soaks using concentrations below 10 mM.

**2-series: chemical elaboration of 2**

Compound **2** contains several sites that were identified for manipulation in structure-activity relationship (SAR) studies ($R_{1\text{-}4}$, Table 4.5). Initially, **3** was synthesized (Kwang Jun Lee, University of Adelaide), in which the phenol was replaced with a methoxy group at position $R_1$. This did not noticeably affect affinity ($K_D$ = 3.4 ± 1.3 mM; Table 4.5), and in fact the increase in molecular mass resulted in a minor decrease in ligand efficiency (0.18 kcal/mol/heavy atom). Accordingly, removal of this moiety completely (**4**) led to increased ligand efficiency (0.20 kcal/mol/heavy atom; molecule synthesized by Birgit Gaiser, University of Adelaide) despite a negligible affinity difference ($K_D$ = 3.5 ± 1.4 mM; Table 4.5). A crystal structure of compound **4** in complex with *Mt*DTBS revealed a similar binding pose to **2**, with slight variation in one chain of the crystal structure (both poses in Figure 4.7A). In Chain D, the electron density indicated that the ketone of **4** ($R_4$) was rotated compared to **2**, suggesting that this group has more conformational freedom in the absence of the $R_1$ phenol group. It is therefore possible that intramolecular hydrogen bonding between these groups may rigidify the structure of **2**. Whether or not this has an effect on the entropy of binding remains a future direction of this project; investigation of the optimal substituent here will require additional studies using higher affinity starting compounds.

**Table 4.5:** SPR determined affinity and ligand efficiencies of **2** and analogues. LE = -2.303(RT×Log($K_D$))/HAC (Hopkins *et al*., 2014).




| Compound | R1 | R2 | R3 | R4 | $K_D$ by SPR (μM) | Number replicates | Ligand efficiency (Kcal/mol/heavy atom) |
|---|---|---|---|---|---|---|---|
| **2** | OH | CCOOH | H | C=O | 3400 ± 400 | 4 | 0.19 |
| **3** | OMe | CCOOH | H | C=O | 3400 ± 1300 | 2 | 0.18 |
| **3a** | OMe | $C(COOH)_2$ | H | C=O | 46 ± 2 | 3 | 0.27 |
| **4** | H | CCOOH | H | C=O | 3500 ± 1400 | 3 | 0.20 |
| **5** | H | CCOOH | COOH | C=O | 19 ± 5 | 3 | 0.32 |
| **6** | H | CCOOH | CCOOH | C=O | 17 ± 1 | 4 | 0.31 |
| **7** | H | CCOOH | COOH | H | 39 ± 2 | 4 | 0.29 |
| **8** | H | CCOOH | CCOOH | H | 56 ± 8 | 3 | 0.29 |

Compound **3a**, an intermediate from the synthesis of **3**, was also tested for its affinity to *Mt*DTBS. This compound contained an additional carboxyl group at position $R_2$, affording it a 74-fold improved affinity over **3** ($K_D$ = 46 ± 2 μM), and an improved ligand efficiency (0.27 kcal/mol/heavy atom; Table 4.5). Crystallographic investigation of the **3a**-*Mt*DTBS interaction revealed it adopted an analogous binding pose to **2**, but with the additional carboxyl group hydrogen bonding with the side chains of Lys37 and Asn40, providing an explanation for the increase in affinity (3.0 – 3.3Å; Figure 4.7B). Lys37, in particular, is a residue involved in stabilizing the phosphorylated DAPA-carbamate, indicating that this substituent also mimics the reaction intermediate. However, the additional carboxyl was also proximal to the side chain of Asp49 (2.6 – 3.5Å), a residue known to coordinate $Mg^{2+}$ during NTP binding (Thompson et al., 2018a) (Chapters 2, 3). This was an unexpected result as two negatively charged carboxyl groups would have been predicted to repel each other, and polar contacts were not predicted by MacPyMOL (Schrödinger). It was therefore hypothesized that these groups would coordinate $Mg^{2+}$, which was present in the SPR running buffer during affinity measurement. However, when $Mg^{2+}$ was included in the soaking experiment, this ion was not seen to coordinate **3a**. While this result demonstrates that growing **2** towards the P-loop region is a viable strategy to increase potency, the crystallographic data suggest that more optimal substituents at the $R_2$ position may be employed to favourably interact with the side chain of Asp49.

The $R_3$ position on the aromatic ring of **2** and **3** superimposes well with C3 of the DAPA alkyl chain (Figure 4.6A, B, Figure 4.7C). This suggested that addition of a carboxyl group at $R_3$ would mimic the interactions made between the DAPA carboxyl and residues Gly144, Leu146 and Asn147 on α6 of the adjacent *Mt*DTBS subunit (2.7 – 3.0 Å; Figure 4.7C). The importance of this interaction site in *Mt*DTBS is supported by the observation that $SO_4^{2-}$ bound to this region in the **2**-*Mt*DTBS structure, suggesting a propensity for binding to anions. Gratifyingly, incorporation of a carboxy (**5**; $K_D$ = 19 ± 5 μM) or carboxymethyl (**6**; $K_D$ = 17 ± 1 μM) group at the $R_3$ position of **3** afforded a 184-fold increase in affinity for *Mt*DTBS, and significantly improved ligand efficiency (Table 4.5). Minimal affinity difference was observed between a one or two carbon carboxyl substituent, however the smaller chain had marginally higher ligand efficiency (0.32 vs 0.31, Table 4.5). Crystal structures of *Mt*DTBS bound to **5** or **6** each confirmed that these carboxy substituents mimicked the aforementioned DAPA interactions (2.7 – 3.4 Å; Figure 4.7D). Minimal structural differences were observed between **5** and **6**, with the latter compound shifting slightly in its pose to accommodate the longer carboxymethyl moiety (Figure 4.7E).

Another pair of carboxy (**7;** $K_D$ = 39 ± 1 μM) and carboxymethyl (**8**; $K_D$ = 56 ± 8 μM) analogues was produced, this time lacking the $R_4$ ketone functionality. These compounds demonstrated that the ketone functional group provides at least a 2-fold increase in affinity for *Mt*DTBS, while also providing additional confirmation that the $R_3$ carboxy substitutions are important for binding. Crystal structures of **7** and **8** bound to *Mt*DTBS reveal that the absence of the ketone does not alter the binding pose compared to molecules **5** and **6** (**5** and **7** shown in Figure 4.7F). In the crystal structure of **5** bound to *Mt*DTBS, this ketone group forms a hydrogen bond to a water molecule (3.28 Å; chain A) that in turn hydrogen bonds to the $R_2$ carboxy group of the compound (2.51 Å), as well as other ordered water molecules within the active site. It is possible that this interaction contributes to the increase in affinity observed in the presence of the ketone moiety. Altogether, these compounds constitute the successful replication of the interactions formed by DAPA within the binding pocket, as well as preliminary SAR of the $R_3$ and $R_4$ positions.

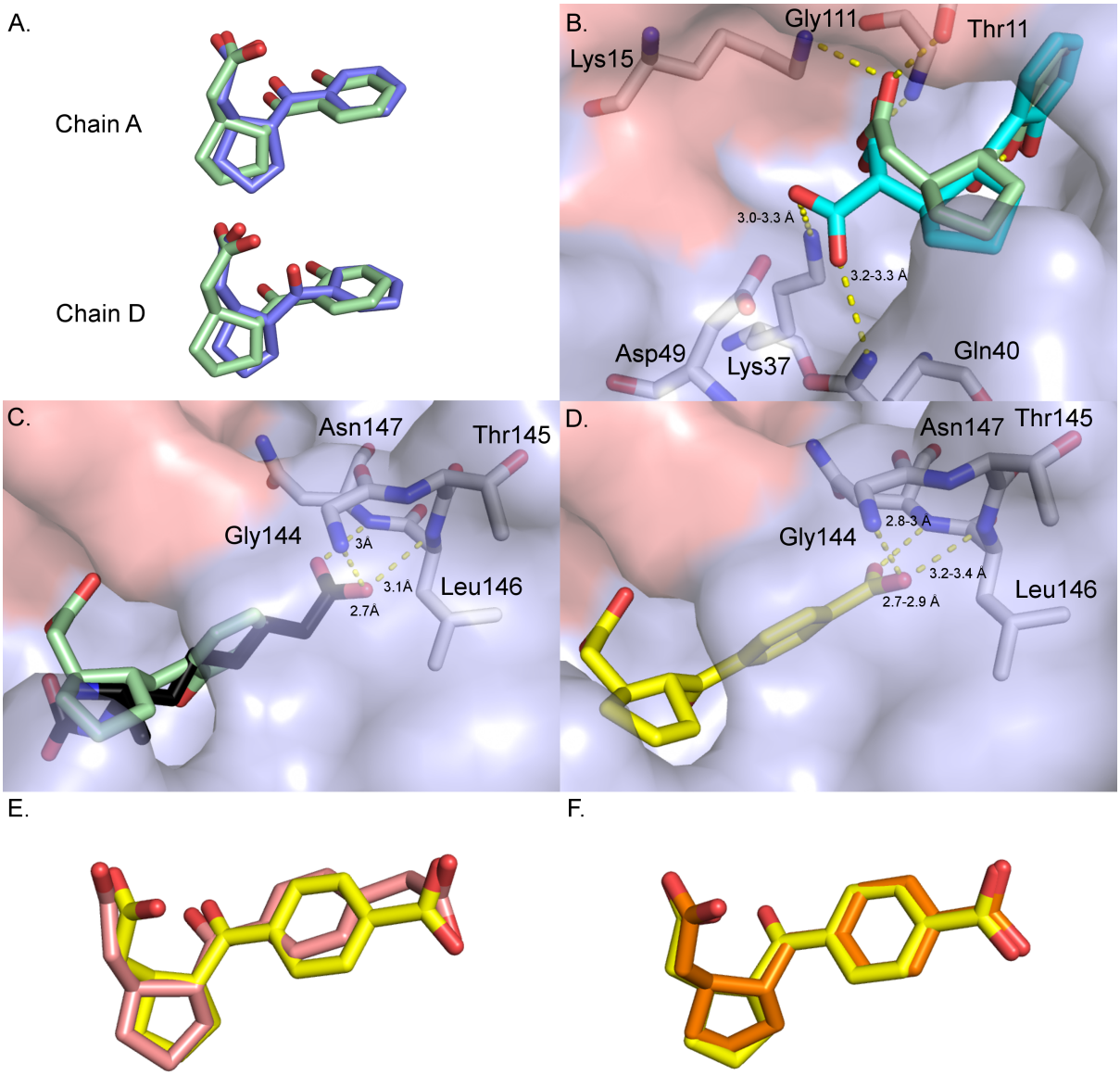



**Figure 4.7:** X-ray crystallography during development of the **2**-series. **A.** Overlay of **2** and **3** (green and blue sticks, respectively) reveals similar binding poses (top), except for chain D (bottom), where the ketone of **3** is rotated compared to **2**. **B.** Binding pose of **3a** (cyan sticks) overlaid with **2** (green sticks) in the *Mt*DTBS active site (blue surface). **3a** binds to the same P-loop residues (salmon surface, grey sticks) as **2**, but the additional carboxyl group faces towards Asp49. The two carboxyl groups are arranged perpendicularly. **C.** The benzyl group of **2** (green sticks) overlays well with C3 of DAPA (black sticks) within the binding pocket (blue surface). DAPA forms hydrogen bonds to residues Gly144, Leu146 and Asn147 on α6 the adjacent subunit (α-helices depicted in Figure 4.2). **D.** Compound **5** (yellow sticks) mimics the DAPA binding pocket interactions previously described. **E.** Compounds **5** (yellow sticks) and **6** (pink sticks) have analogous binding poses, however the carboxymethyl group of **6** induces a slight conformational change compared to **5**. **F.** Compound **7** (orange sticks) does not contain the $R_4$ ketone moiety, however it still adopts an essentially identical binding pose to **5** (yellow sticks). NB: in all cases, only the RS stereochemistry is shown for ease of interpretation. At the time of writing, the preferred stereochemistry of these compounds is unknown.

**Testing 2-series compounds for inhibition of *Mycobacterial* growth**

Select compounds from the **2**-series were tested for antibiotic activity against a *Mycobacterium tuberculosis* cell culture (Figure 4.8; performed by Gayathri Nagalingam, Laboratory of A/Prof Jamie Triccas, Department of Infectious Diseases and Immunology, University of Sydney). Here, *Mtb* growth in a biotin depleted media was measured in the presence of a range of compound concentrations. Despite the low-micromolar affinity of compounds **3a** and **5-8** for *Mt*DTBS, none inhibited the growth of *M.tb* strain H37Rv at concentrations up to 100 μM. It is unknown whether the compounds were able to enter the bacterium. However, it is likely that compounds with higher *Mt*DTBS affinity will be required to inhibit *Mtb*. Considering the $R_2$ carboxyl group of **2**-series compounds occupies a similar region to the γ-phosphate of CTP, these ligands would likely compete with nucleoside triphosphates for binding to *Mt*DTBS. As CTP exhibits a 100-fold higher affinity over the highest affinity compound in the **2**-series (CTP $K_D$ = 160 nM), this ligand would likely outcompete any current **2**-series compound bound to *Mt*DTBS. This could be confirmed in the future, by determining enzyme inhibition by this compound series using the previously established enzyme activity assay (Salaemae *et al*., 2015).

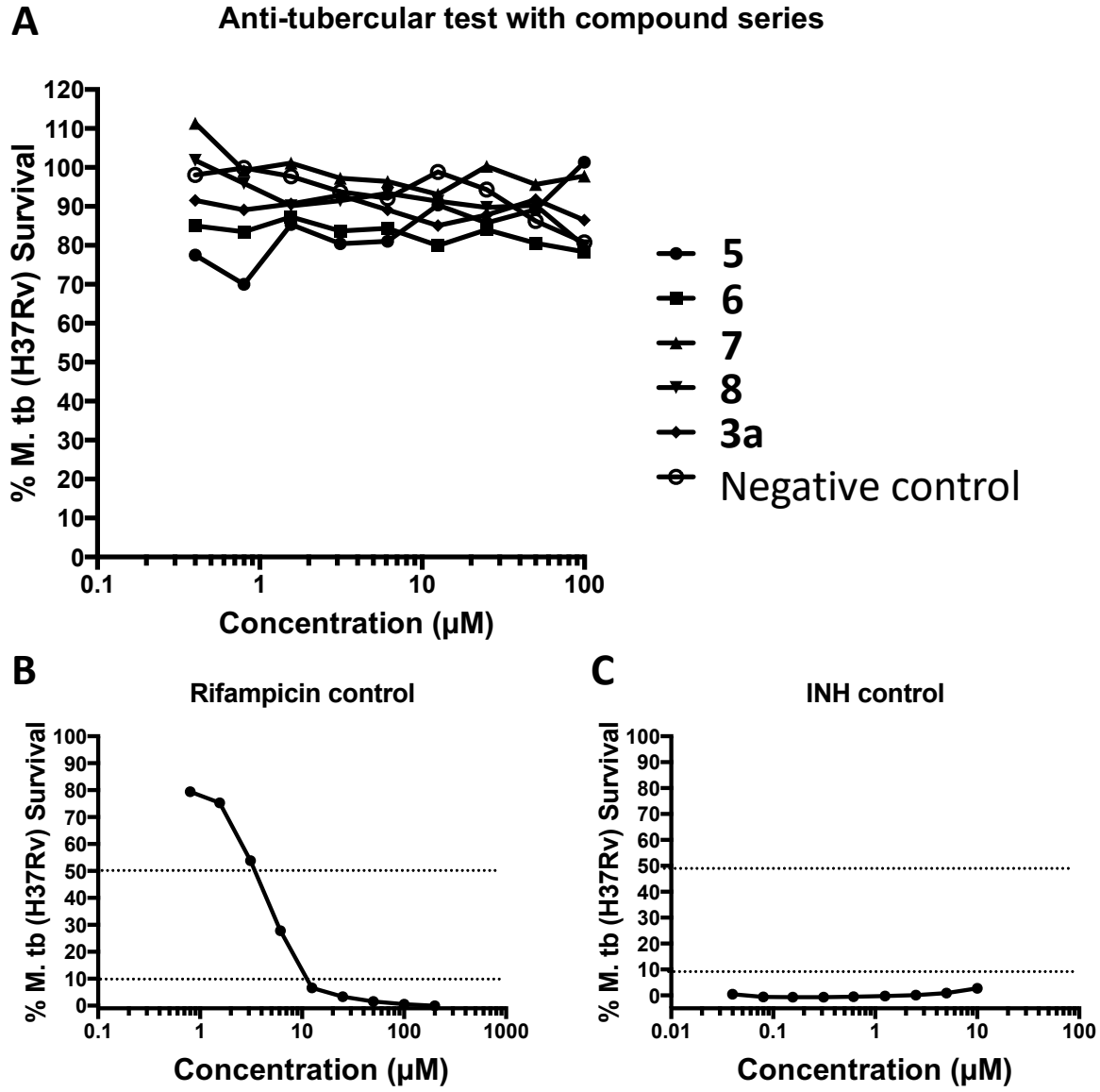



**Figure 4.8:** Testing the antibiotic effect of the **2**-series against *Mycobacterium tuberculosis* strain H37Rv. Experimental work performed by Gayathri Nagalingam, Laboratory of A/Prof Jamie Triccas, Department of Infectious Diseases and Immunology, University of Sydney. **A.** Compounds **5-8** and **3a** did not significantly inhibit the growth of the bacterium. An undisclosed synthetic intermediate of **2** was included as a negative control. **B, C.** Positive controls with the antitubercular drugs rifampicin and isoniazid (INH) demonstrated inhibition of *Mtb*.

### 4.3.4 Linked molecule development and optimisation

**_Mt_DTBS is able to simultaneously bind both cytidine and 2**

Having identified optimised ligands for the cytidine and DAPA binding pockets, compound development efforts turned to the generation of linked molecules. Initially, as proof of concept, a crystal structure of *Mt*DTBS in complex with both cytidine (CTN) and **2** was solved to 2.46Å. Both molecules retained their previously determined binding poses, and $SO_4^{2-}$ was bound to the LRLR nest within the P-loop, as found in previous structures (Figure 4.10). This structure demonstrates that the *Mt*DTBS active site is capable of simultaneously accommodating these compounds, suggesting that it may also be able to bind a compound consisting of CTN and **2** linked together. Consequently, efforts turned towards the development of an appropriate linker moiety for the conjugation of deoxycytidine and an analogue of **2**, to produce a ligand with super-additive affinity (Section 4.1).

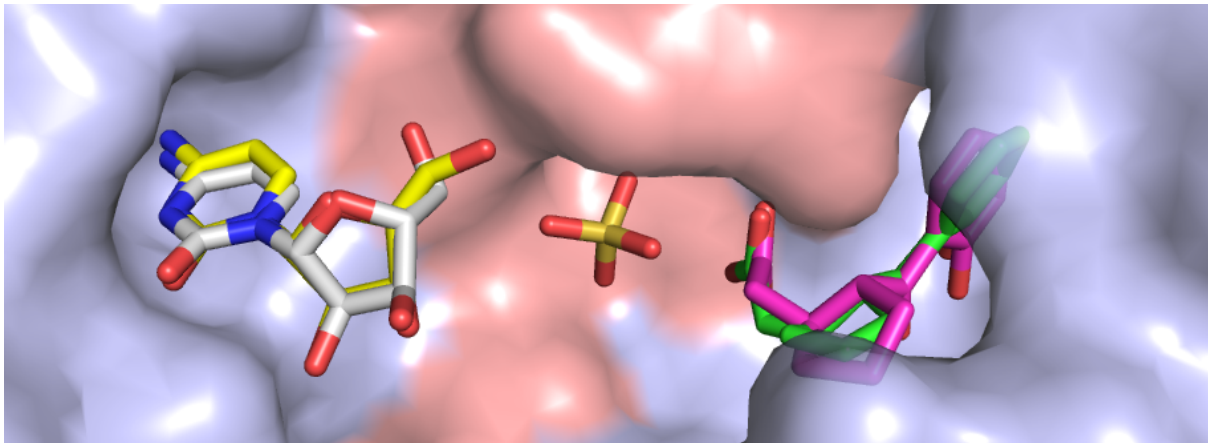


**Figure 4.10**: A soak with both CTN (yellow sticks) and **2** (green sticks) yielded a crystal structure with these compounds bound in identical poses to those previously determined (CTN: grey sticks; **2**: magenta sticks). This revealed that the *Mt*DTBS binding site (blue surface) is able to accommodate both of these ligands simultaneously. $SO_4^{2-}$ was observed bound to the P-loop (salmon surface), as previously shown (PDB ID: 4WOP; Chapters 3, 4).

**First and second generation linked molecules did not bind to the P-loop**

A first generation linked compound was synthesized using Cu-catalysed azide-alkyne click chemistry: an azide derivative of dCTN and an alkyne containing analogue of **2** were reacted to produce triazole-linked compound **9** (synthesized by Kwang Jun Lee, University of Adelaide; structure in Table 4.3). While SPR binding assays with this compound suggested higher affinity than the original substituents, a projection of the $K_D$ beyond the range of concentrations used in the assay was required due to solubility issues with the compound ($K_D$ = 600 ± 200 µM, n = 2, sensorgram in appendix 2; Compare to dCTN $K_D$ = >1 mM, **2** $K_D$ = 3.4 mM). Consequently, this value could be subject to larger error than indicated, and conclusions from this should be cautious. This compound exhibited poor ligand efficiency, presumably owing to its high heavy atom count compared to its substituent compounds (39 heavy atoms; LE = 0.11 kcal/mol/heavy atom). Conventional X-ray crystallographic techniques were unsuccessful in generating ligand complexed crystal structures, however, precipitant-ligand exchange with molecule **9** (crystallographic details in Materials and Methods) yielded a structure that contained reduced biased electron density in the active site (Polder, 3σ) that closely resembled the ligand (Figure 4.11A, B; 0.75 occupancy; 0.87 PHENIX real-space correlation coefficient). Compound **9** was modelled using RS-**2** stereochemistry at the cyclopentane ring. While the dCTN moiety of this molecule bound in

the same pose observed previously, the **2**-moiety differed slightly, presumably due to conversion of the $R_4$ carboxyl group to an amide, abrogating several interactions with the *Mt*DTBS active site and allowing accommodation of the linker moiety (Figure 4.11C, D; compare to Figure 4.6C). Interestingly, the triazole linker moiety did not bind to the P-loop, and instead was located away from the protein surface, around a $SO_4^{2-}$ ion bound at this position (Figure 4.11A). The carbonyl from the amide linkage was proximal to the side chain of Thr11 (2.72Å) and the $SO_4^{2-}$ (2.97Å), suggestive of intermolecular hydrogen bonding or electrostatic interactions. However, the electron density representing the linker moiety is less defined than that of the dCTN and **2** moieties (Figure 4.11B), suggesting that this region bound more dynamically. Consequently, it is possible that these observed hydrogen bonds are more transient. Nevertheless, the SPR and crystallography results suggest that the triazole linker moiety does not favourably interact with the P-loop. This could be a result of several structural features of the ligand, including but not limited to the length and chemical properties of the linker region.

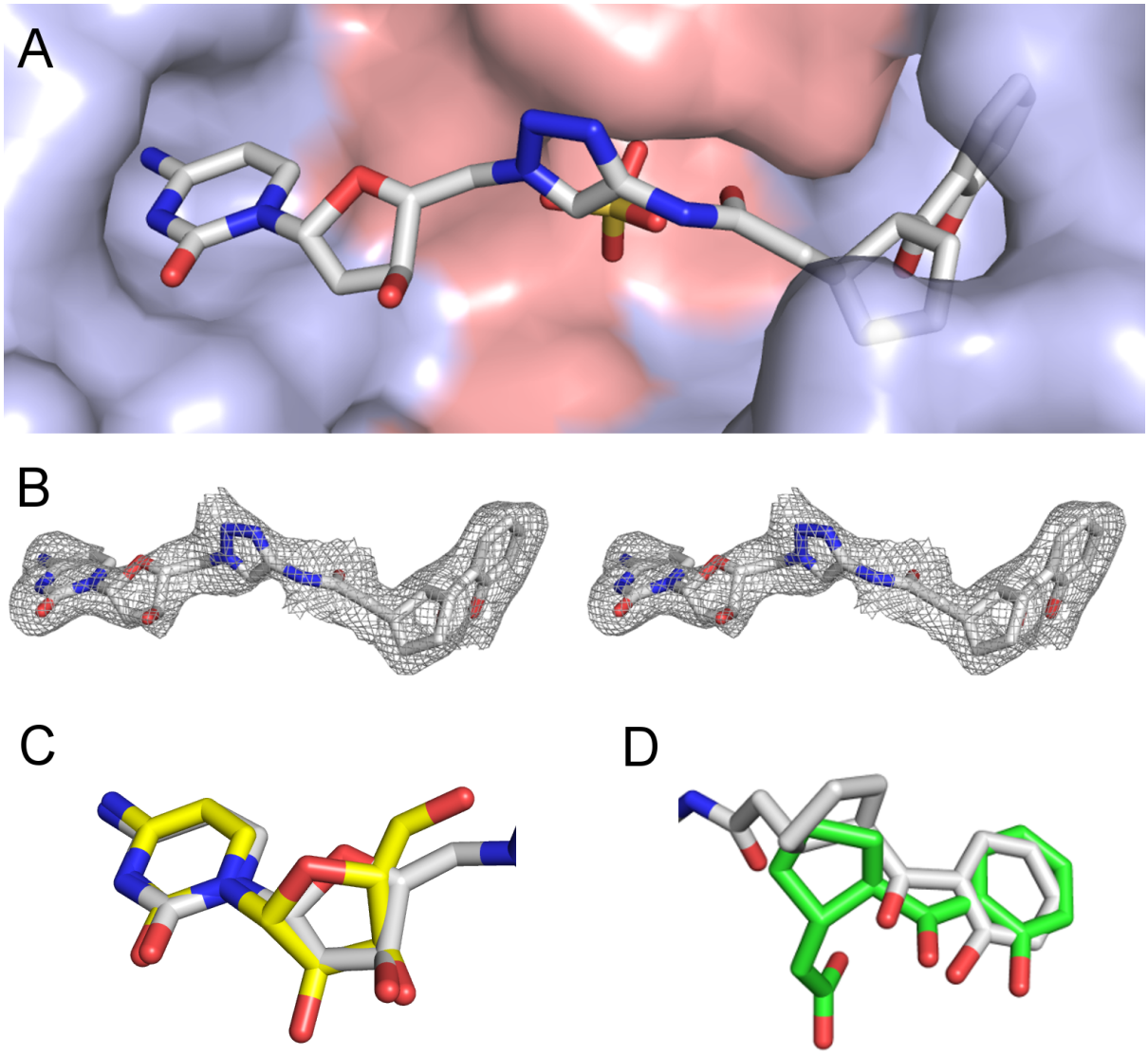



**Figure 4.11:** The binding mode of linked compound **9** to *Mt*DTBS. **A. 9** (grey sticks) binds to the *Mt*DTBS active site (blue surface) above a sulfate ion (yellow sticks) that is bound to the P-loop (salmon surface). **B.** Compound **9** modelled into polder map electron density at 3σ. **C.** The dCTN moiety of **9** (grey sticks) overlays well with CTN (yellow sticks). **D.** The **2**-moiety of **9** exhibits conformational differences from the original **2** binding pose (green sticks), presumably due to the replacement of the $R_2$ carboxyl group to form the amide bond to the linker. In compound **2**, this carboxyl group formed hydrogen bonds with residues Thr11, Lys15 and Gly111 in the *Mt*DTBS active site.

The SPR and X-ray diffraction data suggest that the neutral triazole moiety did not interact with the LRLR nest within the P-loop, which has been shown to have a binding preference for anionic compounds (Chapters 2, 3). A second generation linked compound **10** was synthesized containing a glutamic acid linker moiety (Kwang Jun Lee, University of Adelaide; structure in Table 4.3). This compound contained dCTN conjugated to the glutamic acid side-chain, and **6** conjugated to its amine group via an amide bond. The free carboxyl moiety of the glutamic acid linker was hypothesised to bind to the LRLR nest within the P-loop. When combined with the presence of a carboxymethyl group at the $R_4$ position of **6**, this compound was expected to have higher affinity for *Mt*DTBS than triazole linked compound **9**. However, SPR binding experiments with **10** revealed the binding affinity and ligand efficiency of the two compounds were similarly low ($K_D$ = 700 ± 200 μM vs 600 ± 200 μM for **9**; LE = 0.10 kcal/mol/heavy atom vs 0.11 kcal/mol/heavy atom for **9**). The affinity of **10** was ~40-fold weaker than **6** itself. X-ray crystallographic investigation by traditional soaking methods revealed electron density that displayed a similar binding pose to **9**, wherein the dCTN and **6** moieties adopted poses similar to those previously observed, but the linker moiety bound adjacent to the crystallographic precipitant, citrate (Figure 4.12A-D; 0.69 occupancy; 0.862 PHENIX real-space correlation coefficient). Here, electron density for the glutamic acid linker moiety was weaker than that for the dCTN and **6** substituents, indicating that it did not possess a single, stable conformation but was instead disordered (Figure 4.12B). The SPR and X-ray crystallographic data show that using glutamic acid as the anionic linker moiety to conjugate dCTN and **6** did not improve binding over the triazole linked compounds. Considering that both of these linked molecules are able to bind adjacent to the crystallographic precipitant at the P-loop, the length of the linker moiety may need to be reduced to encourage interaction between the linker and the P-loop.

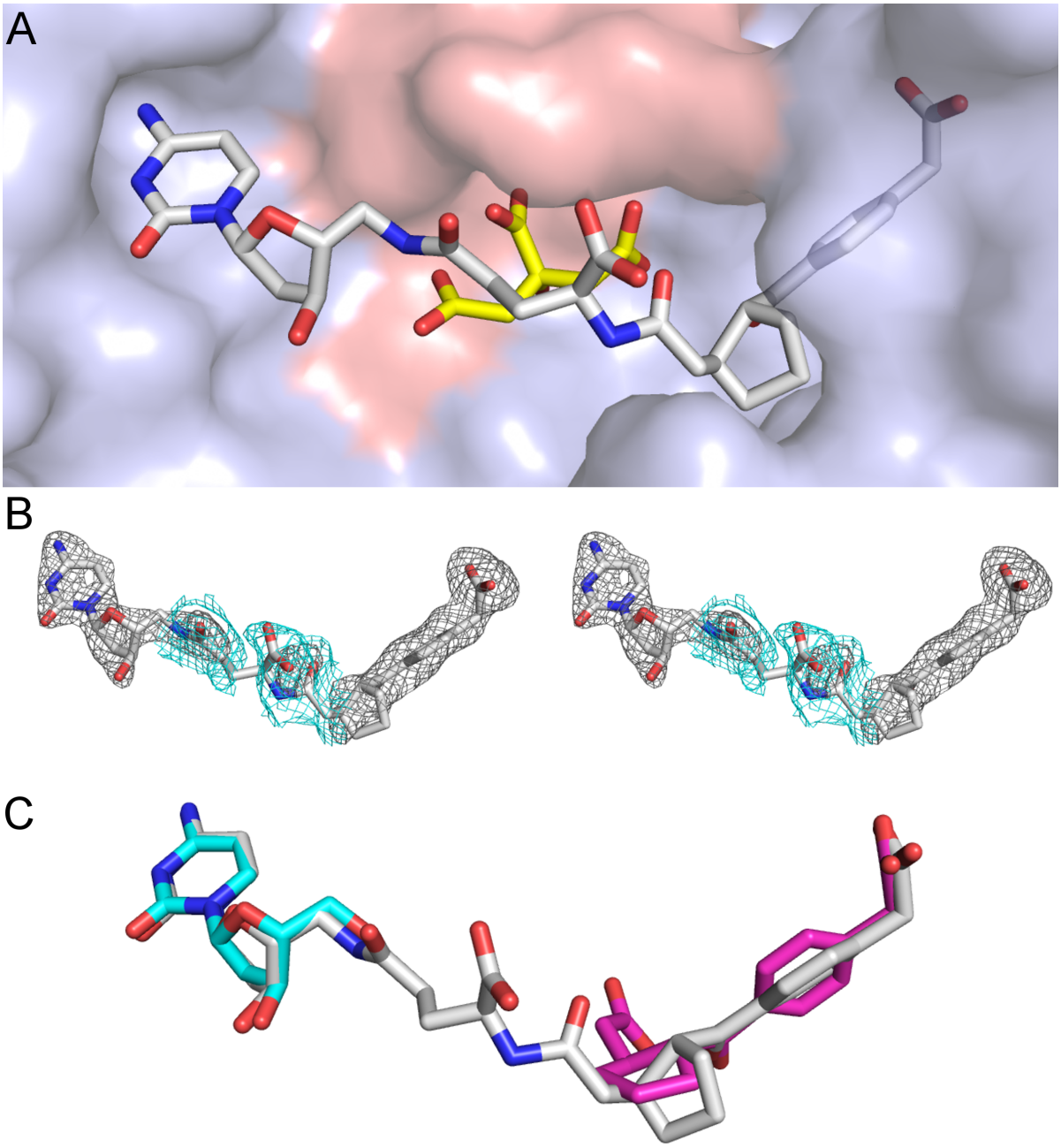



**Figure 4.12:** The binding pose of **10** (grey sticks) in the *Mt*DTBS active site (blue and salmon surface). **A.** Like **9**, **10** exhibits a binding pose adjacent to the crystallographic precipitant, in this case citrate (yellow sticks), but the dCTN and **6** moieties bind to the cytidine and DAPA binding pockets respectively. **B.** Polder electron density for **10** is strong (3σ; grey mesh) for dCTN and **6** but weak (1σ; cyan mesh) for the glutamic acid linker moiety, suggesting it has a less stable conformation, and may bind in a more disordered manner. **C.** Overlay of **10** with crystal structures of dCTN (cyan sticks) and **6** (magenta sticks) individually. The dCTN and **6** substituents of **10** bind to similar regions to these individual compounds. Similar to compound **9**, the **6** substituent on **10** differs in its pose, presumably due to the modification of the $R_2$ carboxyl group during conjugation to the glutamic acid linker. In compound **6**, this carboxyl group hydrogen bonds to Thr11, Lys15 and Gly111.

## 4.4 Discussion and Future Directions

### 4.4.1 Investigation into the cytidine binding pocket

The crystallographic screening of the cytidine analogues revealed several key points regarding the nature of the cytidine binding site. Firstly, the absence of structures of *Mt*DTBS bound to alternate nucleosides and nucleoside analogues containing varied base moieties provides additional confirmation that the cytidine binding site has evolved to be highly specific for the cytosine base. However, the use of crystal soaking and X-ray crystallography in this circumstance could bias these observations towards the energy minimized states of the protein structures that form the cytidine binding site. It is possible that, in solution, the cytidine binding site is more flexible and could undergo conformational changes which would allow it to accommodate alternate scaffolds. Experiments to investigate this could include solution-based ligand-protein interaction detection techniques such as heteronuclear single quantum coherence nuclear magnetic resonance (HSQC NMR). Here, *Mt*DTBS would be labelled with $^{15}N$, allowing for the observation of the chemical environment of each nitrogen atom within the enzyme. Spectra would be recorded in the absence and presence of cytidine, generating a pattern of residues that it interacts with in the binding site. By comparing these spectra to those determined in the presence of alternate nucleosides and nucleoside analogues, it would be possible to ascertain whether or not such compounds bind to the cytidine binding site residues. A concern with the use of HSQC NMR is that spectrum quality is generally inversely proportional to protein molecular weight. The *Mt*DTBS homodimer has a molecular weight of 47 kDa in solution, which could prove problematic for this technique. Initial attempts to obtain HSQC NMR spectra did reveal observable and distinguishable peaks (data not shown). However, optimisation efforts were focussed on SPR and X-ray crystallography in the current project. Nevertheless, this does indicate that optimisation of HSQC NMR to study *Mt*DTBS in future projects is feasible.

Crystallographic screening also determined that the cytidine binding pocket is able to accommodate structural differences at the 2′ position on the ribose ring of cytosine containing nucleosides. Crystal structures were determined with dCTN and dFdC, containing unsubstituted 2′ and 2′-difluoro moieties respectively, compared to the 2′-hydroxy group of CTN. The ability to accommodate variations at the ribose ring may be due to the relatively solvent exposed nature of the *Mt*DTBS cytidine binding site, as discussed in Chapter 3. The 2′ ribose position is located adjacent to hydrophobic residues, which may interact unfavourably with the hydrophilic 2′-hydroxy moiety. This provides a possible explanation for the decreased association rate and lower affinity observed for 2′-hydroxy containing nucleotides. While the effect of the 2′-difluoro substitution was not determined, it could be investigated in the future through SPR binding affinity assays with the pharmacologically active triphosphate form (dFdCTP) (Alvarellos *et al*., 2014). It is hypothesized that, due to the hydrophobic nature of C-F bonds, the 2′-difluoro substituent may also interact favourably with the hydrophobic residues. Comparison to dCTP and CTP would be useful for future inhibitor development. Further exploration of the 2′ ribose site, as well as the ribose binding site as a whole, may reveal alternate functional groups that result in even higher binding affinity. A final consideration is that incorporation of fluorines in therapeutics has a number of pharmacological benefits such as improved metabolic stability and membrane permeability that could warrant inclusion of the 2′-difluoro, regardless of its affinity effects (Shah and Westwell, 2007).

By assaying the di- and triphosphate counterparts of dCTN with SPR, it was determined that proton substituents at the 2′ site were preferred by *Mt*DTBS over the hydroxy group. This finding invites further investigation into *Mt*DTBS NTP utilization in a biological setting, as it is possible that dCTP is the catalytically preferred substrate *in vivo*. The likelihood of this depends on the concentrations of the nucleoside triphosphate pools available to *Mt*DTBS at varying stages of the *Mtb* lifecycle, a factor that is currently unknown, as well as confirmation of the presumed utility of dCTP in the enzyme reaction. The latter could be determined using the previously established *Mt*DTBS enzyme activity assay (Salaemae *et al.*, 2015). Performing this assay in the presence of a mixture of NTPs at biologically relevant relative concentrations and observing the relative rates of substrate utilization (e.g. by LC/MS and or NMR) would reveal valuable insight into the *in vivo* substrate promiscuity of *Mt*DTBS.

### 4.4.2 *In silico* fragment hit investigation: discovery and SAR of compound **2**

The discovery of degradation product **2** as a ligand of *Mt*DTBS was an improbable and serendipitous circumstance. It highlights the utility of the model provided by X-ray crystallography in the drug discovery process, as it is likely that **2** would not have been discovered without use of this technique to investigate the *in silico* hits. This discovery also highlights the importance of using orthogonal techniques in screening processes; not only to investigate preliminary hits, but also to confirm their molecular structures, a sentiment that is echoed in the literature (Davis and Erlanson, 2013). However, this compound constitutes a single confirmed hit from a screening campaign and its importance should not be overemphasized. Given the ability of the largely hydrophobic *Mt*DTBS DAPA pocket to also accommodate the adenosine moiety of ADP (Chapter 2), the DAPA binding site, and indeed the whole of the active site, should be probed with additional fragment screening campaigns, using techniques such as SPR, thermal shift assay (TSA) or saturation transfer difference NMR (STD-NMR). Identifying alternate chemical scaffolds that bind to various regions of the *Mt*DTBS active site would increase the capacity to merge, grow or link fragments towards efficient, high affinity inhibitors of the enzyme and *M. tuberculosis.*

The utility of **2** lies in the possibility for chemical modification to probe the *Mt*DTBS active site. Initial SAR concerned position $R_1$, where little affinity difference was observed by SPR between proton, hydroxyl and methoxy substituents, primarily due to the weak affinity observed for these compounds. While a proton substituent in this position was utilized for the development studies outlined in this chapter, affinity differentiation between substituents at this position is still required. This could be achieved by assaying the affinity of $R_1$ variants on a higher affinity scaffold, such as **5** or **6**, akin to the SAR carried out at the $R_4$ position. It is hypothesized that the hydroxyl $R_1$ substituent may stabilize the compound through intramolecular hydrogen bonding with the $R_4$ ketone, increasing the affinity by reducing the entropic penalty upon binding (Andrews *et al.*, 1984; Chang *et al.*, 2007). Similarly, the preferred stereochemistry of the **2**-series ligands remains unknown. This could be determined by SPR binding assays with each stereoisomer (SS/RR/SR/RS) using higher affinity analogues that are able to be physically separated. When assaying higher affinity compounds the results become more precise, leading to an increased ability to confidently distinguish affinities. In doing so, SAR of the **2**-series compounds can be performed more rigorously in order to determine the optimal substituents at positions $R_{1\text{-}4}$.

The binding mode of **2** was compared to the biological substrates of *Mt*DTBS to guide further compound development. The addition of carboxy and carboxymethyl substituents to position

$R_3$ to form **5** and **6** successfully replicated interactions that DAPA makes within its binding site, greatly increasing the affinity of these compounds. This was a highly promising result, and definitively demonstrates the potential for structure-guided compound development against the *Mt*DTBS target. This also implies that replication of interactions between anionic species and the LRLR nest is feasible and is likewise predicted to substantially improve binding affinity. As shown in chapters 2 and 3, this region within the P-loop exhibits extremely high affinity for anionic species and replicating such interactions would likely result in a significant increase in the affinity of potential inhibitors. Indeed, the carboxy at position $R_2$ of compound **2** bound to some P-loop residues, as well as those that are involved in the stabilization of the phosphorylated DAPA-carbamate reaction intermediate (observed in PDB ID: 1DAK)(Käck *et al*., 1998). Compound **3a**, with a dicarboxy substituent at $R_2$, was thought to bind to additional P-loop residues but instead exhibited hydrogen bonding with additional residues that stabilized the reaction intermediate. This was another promising result, leading to a large increase in affinity for **3a** over **3**. Mimicking enzyme transition states has been documented as a successful method of drug discovery (Schramm, 2018) and is a method of compound progression that should be considered for the development of inhibitors from the **2**-series.

There are clear opportunities for increasing the affinity of **2**-series molecules to *Mt*DTBS. Firstly, combining the dicarboxy substituent of **3a** (~70-fold improvement over **3**) with the carboxy or carboxymethyl substituents present in **5** and **6** (~200-fold improvement over **2**) could greatly increase affinity for *Mt*DTBS. It is hypothesized that this could result in as large as a 14,000-fold improvement in affinity, leading to a predicted $K_D$ of ~200 nM. Precipitant-ligand exchange revealed that the additional carboxyl group of **3a** was most proximal to the side chain of Asp49 (2.6 – 3.5Å compared to 3.0 – 3.5Å for Lys37 and Asn40). Replacing this carboxyl group with a positively charged substituent such as an amidine or guanidine (Greenhill and Lue, 1993) could possibly increase the affinity of **3a** as a result of interaction to the negatively charged Asp49 side chain, even if interactions to the Lys37 and Asn40 were lost. However, the charged nature of amidine and guanidine groups may lead to undesirable pharmacological properties, and aspartate-binding groups used in BACE1 inhibitors such as aminohydantoin, aminooxazoline, or aminothiazoline may be preferable (Coimbra *et al*., 2018). Ideally, a functional group that is capable of binding to the Asp49 in addition to the Lys37 and Asn40 would be used, resulting in a highly efficient compound. Finally, further fragment growing into the P-loop to interact with the LRLR nest should be considered, as it could potentially lead to extremely large increases in affinity. Currently, chemical modification of **2** has resulted in a ~180-fold increase in affinity, which is a highly promising start to the hit to lead development process and provides a solid foundation for further compound development endeavours.

### 4.4.3 *M. tuberculosis* inhibition

There are several possible explanations for the failure of **2**-series compounds to inhibit *Mtb* growth in a culture based assay. The *Mtb* cell wall consists largely of mycolic acids, a component which is thought to contribute to virulence and to impede cell penetration by drugs, especially those that are hydrophilic (Lambert, 2002; Takayama *et al*., 2005). Rifampicin, a clinically used antibiotic for the treatment of TB and a positive control in the reported assay, is relatively hydrophobic, enabling it to cross the mycolic acid barrier and inhibit the bacterium. Smaller and more hydrophilic TB drugs such as isoniazid (INH in the *Mtb* assay), pyrazinamide and ethambutol are thought to traverse the mycobacterial cell wall through membrane bound porin proteins, which exhibit low rates of uptake (Lambert, 2002;

Masi *et al*., 2017). The most potent compounds from the **2**-series each contain 2 anionic carboxylic acid groups, increasing the polarity and hydrophilicity of these compounds and making traversal of the cell wall barrier most likely to occur through similar porin structures. Considering that the maximum concentration used in the *Mtb* inhibition assay was 100 μM, it is likely that the concentration of **2**-series compound within the cell would be even less. Furthermore, given that the affinity difference between the substrates CTP or dCTP and **6**, the highest affinity **2**-series molecule, is >100-fold, it is possible that these compounds would have been outcompeted for binding to *Mt*DTBS. In order to determine if this is the case, the efficacy of *Mt*DTBS inhibition by **2**-series compounds should be tested in an enzyme activity assay. While providing a possible explanation for the lack of inhibition observed, this also indicates that higher affinity *Mt*DTBS binding compounds will be required in order to achieve successful inhibition of the mycobacterium.

The pharmacological properties of **2**-series compounds should also be modified in order to increase permeability through the mycobacterial cell wall and result in a higher concentration within the cell. This may include increasing lipophilicity in order to allow diffusion through the mycolic acid barrier, as observed with rifampicin. However, factors such as solubility, size, bioavailability as well as other pharmacokinetic parameters should be taken into consideration. One such improvement on the current **2**-series molecules could be to replace existing carboxylic acid groups with more drug-like moieties such as the tetrazole. Tetrazoles have similar pKa values to carboxylic acids but exhibit higher lipophilicity due to the delocalization of electrons around the ring (Myznikov *et al*., 2007). Consequently, tetrazole containing compounds could retain interactions to the positively charged residue backbones such as those seen between **2**-series molecules and the P-loop or the DAPA pocket, while potentially resulting in an increased ability to traverse the mycobacterial cell wall. Indeed, tetrazole containing compounds have been reported in antifungal compounds (Lockhart *et al*., 2016; Warrilow *et al*., 2014). Ideally, tetrazole substitution would retain the binding affinity of **2**-series compounds, whilst increasing their drug-like nature, however, it is possible that interactions will not be fully retained and affinity for *Mt*DTBS may decrease. Future studies hoping to simultaneously optimise affinity and pharmacological through the use of substitutions such as the tetrazole should monitor the affinity and inhibition of both enzyme and mycobacterium for each compound.

### 4.4.4 Linked molecule development

The identification of ligands bound to each pocket of the *Mt*DTBS active site encouraged the development of a linked compound with potentially super-additive affinity. The two synthesized linked compounds **9** and **10** showed the potential of this approach, as the dCTN and **2**-series substituents were found to bind to *Mt*DTBS by X-ray crystallography, adopting similar poses to those previously determined. Given this, it was expected that the compounds would bind with higher affinity than their substituent molecules. While this appeared to be the case for **9**, it was not for **10**. This could be because X-ray crystallography provides a static model, which can be different from what occurs in solution (Davis and Erlanson, 2013). This is particularly pronounced when observing dynamic phenomena. Consequently, the structures with **9** and **10** may not be fully representative of the protein-ligand interaction that is observed in solution by SPR. For example, it may be that low affinity was determined by SPR because the dCTN and **2**-series moieties of the linked compounds do not simultaneously bind to the *Mt*DTBS active site in solution. Additionally, precipitants were observed bound to the P-loop in the crystal structures, but were not present in the SPR experiments to determine affinity. In the crystallography, it is possible that the precipitant molecules aided in providing

order to the *Mt*DTBS active site by binding to the P-loop, or interacted with the linker itself and stabilized the observed binding mode. To test this hypothesis in solution, SPR ligand binding was attempted with precipitants present, however these compounds appeared to bind non-specifically to the sensor surface, presumably due to their charged nature (data not shown). This hindered analysis of the effect that precipitants have on linked compound binding in solution, inviting further investigation by orthogonal methods such as HSQC NMR. Nevertheless, the current crystallographic and SPR results together reveal several options for future development of linked compounds.

Despite occupying a similar region at the DAPA pocket, the binding mode of the **2**-series substituent of the linked molecules differed slightly compared to that observed for compounds **2**-**8**. This was most likely a consequence of abrogation of the hydrogen bonds between the $R_2$ carboxyl group and residues Thr11, Lys15 and Gly111; the interactions that mimic the reaction intermediate demonstrated previously (PDB ID: 1DAK)(Käck *et al*., 1998). In both linked compounds, this group was converted to an amide during conjugation to the linker moiety, and subsequent geometric rearrangement of this region prevented the same hydrogen bonds from being retained. In the future, linkers should aim to retain these hydrogen bonds, possibly by conjugating through one of the two carboxyl groups at position $R_2$ in compound **3a**. This may be difficult due to their structural similarity, and care must be taken to singly protect one carboxyl to ensure that this is done selectively. In order to ensure that linked compounds are highly efficient, the investigation of additional and alternative means to conjugate the linker to the **2**-series compounds will be required.

The composition of the linker moiety is a crucial component to future compound development. The similar affinities observed between compounds **9** and **10** indicate that the triazole linker may in fact be preferable to the glutamic acid linker. Compound **10** was expected to have higher affinity not only due to the linker moiety, but also because of the inclusion of **6**, which exhibited the highest affinity for *Mt*DTBS owing to the presence of the $R_3$ carboxymethyl. By comparison, **9** contained the original compound **2** that does not possess a $R_3$ carboxyl substituent, yet bound with comparable affinity to **10**. This suggests that the glutamic acid linker may negatively affect binding, possibly owing to factors such as geometry, flexibility and strain of the linker (Chung *et al*., 2009). Indeed, the linker moiety of **10** contains multiple rotatable bonds, affording it more conformational flexibility than the more rigid and planar triazole linker of **9**. Thus, there is likely to be a smaller entropic penalty upon binding for **9**, whereas **10** may encounter a large penalty for the loss of rotational and translational entropy (Andrews *et al*., 1984). These factors can be difficult to distinguish using the static crystallographic model, however it is clear that the electron density representing the glutamic acid linker was very weak compared to the rest of the molecule, suggesting that it did not adopt a stable binding pose. This was also reflected in the high B-factors for the glutamic acid linker moiety compared to the rest of **10** and also *Mt*DTBS residues (B-factors: **10**-linker = ~40-55, **10** average = 37, *Mt*DTBS average = 39). In order to determine whether the triazole exhibits more preferable characteristics as a linking moiety, a compound consisting of dCTN linked to **6** by a triazole moiety should be synthesized and tested for binding. However, the neutral triazole linker is not expected to interact with the P-loop. Given that LRLR nest in particular is known to exhibit extremely high affinity for anionic compounds, other linker moieties should be investigated to replicate these interactions. One such example would be the use of an aspartic acid linker, which would reduce the linker by one carbon compared to glutamic acid. This may exhibit different geometry, possibly constraining the linker towards the P-loop, encouraging interactions with the LRLR nest. If successful, this would be expected to result in a large increase in affinity.

Finally, as with the development of the **2**-series, the pharmacological and pharmacokinetic properties of future linked molecules should also be carefully observed and optimised, such that permeability of the *Mtb* cell wall is possible.

### 4.4.5 Concluding remarks

There are countless options for future development of compounds to bind *Mt*DTBS with high affinity. The work presented here constitutes the beginning of the optimisation process, and it is intended that the reported data will be used to guide future compound development. The most useful information provided is the optimisation and use of the methodologies for biophysical and structural investigation of *Mt*DTBS ligand binding. Using this information, synthesized compounds can be rapidly and accurately assessed by orthogonal assays in future studies. This will be an iterative and collaborative process that shows promise for the development of inhibitors of *Mt*DTBS and ultimately *Mtb*.

# Chapter 5:

## Discussion and Future Directions

## 5.1 Antibiotic development through new pathways and techniques

It is clear that significant investment and development is required to combat the rise in antibiotic resistance, as this phenomenon has the potential to become a global human health crisis. In the current project, the biotin biosynthesis enzyme *Mt*DTBS was identified as a promising target for the development of antibiotics to treat drug resistant Tuberculosis. This is a notoriously difficult process, spurring the adoption of innovative methods of drug discovery, as discussed in the introduction (Lamoree and Hubbard, 2018; Tommasi *et al*., 2015). Here, the emerging technique fragment-based lead discovery (FBLD) was employed to identify and develop compounds that bind to *Mt*DTBS with high affinity. In accordance with the ethos that a fundamental understanding of the characteristics of a drug target will aid inhibitor development, the Booker laboratory had already begun to investigate the biochemical nature of *Mt*DTBS catalysis. It was determined that, uniquely, *Mt*DTBS was able to utilize multiple nucleoside triphosphates (NTPs) to catalyze its reaction (Salaemae *et al*., 2015). However, the biophysical and structural mechanisms by which *Mt*DTBS was able to promiscuously bind NTPs were unknown. Consequently, structural and biophysical investigation of these mechanisms constituted a major part of this thesis. Throughout the investigation of this mechanism it was observed that X-ray crystallography of *Mt*DTBS in complex with ligands was hindered by the presence of crystallographic precipitants within the active site. Precipitant-ligand exchange was developed to overcome this issue, and was deemed of use to the crystallographic community (published in *Acta Crystallographica Section D*). Using this technique to determine structures of *Mt*DTBS in complex with NTPs revealed insight into the promiscuity mechanism that was of interest to the catalysis field (published in *ACS Catalysis*). Together, these articles contributed to understanding the ligand binding characteristics of *Mt*DTBS, and aided the decision-making process in regard to the FBLD compound identification and development presented in Chapter 4. It is hypothesized that this information, particularly that regarding the characteristics of the phosphate binding loop (P-loop) will be useful in the future development of inhibitors against *Mt*DTBS. In this discussion, precipitant-ligand exchange will be discussed initially, as it was an enabling technique for many of the results in this thesis. Following this, the structural and biophysical characterisation of *Mt*DTBS ligand binding will be summarised and discussed with regard to the literature, followed by discussion regarding the use of this information for guiding compound development.

## 5.2 Precipitant-ligand exchange is an enabling technology

During the structural investigation of *Mt*DTBS, it became apparent that the P-loop was binding to the anionic crystallographic precipitants sulfate and citrate. This issue was exacerbated by the high precipitant concentration used in the crystallization condition (1.0 - 1.5 M sulfate; 0.7 M citrate). By transferring crystals into a solution that contained the ligand of interest in place of the precipitant, several key ligand-complexed crystal structures of *Mt*DTBS were solved. The development of this technique was crucial to many of the results presented in this thesis. It is hypothesized that this technique will also be of use to the wider crystallography community. X-ray crystallography is the most widely used structural biology technique, contributing 130,743 of the 146,093 (89%) structures deposited in the Protein Data Bank (PDB) at the time of writing (November 2018). Of the X-ray crystallography structures deposited in the PDB, 18,089 (14%) contain a sulfate ion, constituting the most commonly modelled free ligand in the PDB. Incidentally, the P-loop, which bound the sulfate ion, was a major focal point for both the investigation of the *Mt*DTBS NTP binding characteristics, as well as for the future of compound development. It is therefore possible that crystallographic

precipitants have prevented the atomic resolution investigation of other important interactions observed in proteins throughout the proteome. Consequently, precipitant-ligand exchange is intended to be of use to crystallographic researchers throughout the world.

## 5.3 Structural and biophysical characterisation of *Mt*DTBS natural ligand binding

### 5.3.1 *Mt*DTBS exhibits a unique mechanism of NTP promiscuity in catalysis

Here, *Mt*DTBS NTP promiscuity will be briefly summarised, followed by a discussion of this finding in regard to the wider literature. Prior to the work presented in this thesis, *Mt*DTBS was known to promiscuously utilize nucleoside triphosphates (NTPs) for catalysis. However, the mechanism was unknown. In Chapter 3, surface plasmon resonance (SPR) and X-ray crystallography were used to determine that *Mt*DTBS alternates between two NTP binding modes. Cytidine triphosphate (CTP) binds in a high affinity binding mode, forming a complex that dissociates slowly as observed by SPR. X-ray crystallographic structural data for this interaction revealed strong electron density representing the entirety of the ligand, and hydrogen bonding was observed between the cytosine moiety and the enzyme. All other NTPs tested against *Mt*DTBS by SPR revealed fast association and dissociation rates, with similar affinities regardless of the varying nucleoside. X-ray crystallographic investigation of these NTPs exhibited significantly stronger electron density for the triphosphate moiety than the nucleoside. This lead to the conclusion that ligands binding in the promiscuous binding mode were predominantly 'anchored' to *Mt*DTBS through the interaction between the P-loop and the triphosphate, while the nucleoside moiety was less ordered. This binding mode was visualized through the use of molecular dynamics simulations. Finally, the binding affinity contribution of the triphosphate moiety was quantified, with the interaction between the β-phosphate and a moiety within the P-loop known as the LRLR-nest exhibiting extremely high affinity. Together, these results demonstrate that *Mt*DTBS can promiscuously utilize NTPs through dual binding modes, both of which are governed by similar, high affinity interactions between the P-loop and triphosphate tail.

The mechanism of NTP promiscuity observed for *Mt*DTBS has not been observed in homologous enzymes from other bacteria. In structurally characterized DTBS enzymes, differences in the nucleoside binding pockets and P-loops may provide possible explanations as to the unique nature of the *Mt*DTBS catalytic mechanism. These were briefly summarised in chapter 3, and will be discussed in more detail and with regard to available biophysical and biochemical data. We have previously demonstrated that *Escherichia coli* DTBS (*Ec*DTBS) has a more than 100-fold catalytic preference for ATP over CTP and GTP, and cannot use ITP, UTP or TTP to carry out its reaction (Salaemae *et al*., 2015). Structures of *Ec*DTBS bound to ATP reveal electron density corresponding to the entire ATP molecule (PDB ID: 1A82), implying a single, stable conformation akin to CTP-MtDTBS. This conformation involved hydrogen bonding between adenine and the sidechain of Asn175. The spatially equivalent amino acid in *Mt*DTBS is a glycine, which is unable to support the same side chain interaction. Crystal structures have also been determined for the *Helicobacter pylori* DTBS (*Hp*DTBS) in complex with both ATP (PDB ID: 3QXC) and GTP (PDB ID: 3QXX), although binding affinities for these ligands were not reported (Porebski *et al*., 2012). The *Hp*DTBS structures also reveal interactions between the nucleoside and Asn175, but different contacts are made between the two purines and *Hp*DTBS. This method of promiscuity is clearly distinct from the mechanism described for *Mt*DTBS, where interactions between promiscuous nucleosides and the enzyme were not observed. Furthermore, the nucleoside electron density in these other DTBS structures was well-defined, indicating that binding

occurs in a single conformation. The absence of a hydrogen bonding partner in *Mt*DTBS equivalent to Asn175 is a plausible explanation as to why *Mt*DTBS was not observed to bind the nucleoside of ATP or other NTPs. An alternative explanation arises from the observation that *Ec*DTBS also binds the ATP ribose through hydrogen bonding with residue Glu211, another residue that is not present in *Mt*DTBS. The absence of hydrogen bonding between DTBS and the nucleoside ribose sugar may allow for the nucleoside conformational flexibility that was observed in the ensemble refinement of ATP-*Mt*DTBS (chapter 3). Interestingly, an equivalent of the hydrogen bond between *Ec*DTBS Glu211 and the nucleotide ribose is not present in *Hp*DTBS. Assessment of *Hp*DTBS nucleotide specificity would therefore be useful for clarifying the role of this interaction. An additional feature that may contribute to the promiscuity mechanism of MtDTBS is sequence variation in the P-loop. The *Mt*DTBS (Gly-Thr-Gly-Thr-Gly-Val-**Gly-Lys-Thr**), *Ec*DTBS (Gly-Thr-Asp-Thr-Glu-Val-**Gly-Lys-Thr**), *Hp*DTBS (Ala-Thr-Asn-Thr-Asn-Ala-**Gly-Lys-Thr**) and *Francisella tularensa* DTBS (Gly-Thr-Asp-Thr-Glu-Val-**Gly-Lys-Thr**) all share a Gly-Lys-Thr motif (bold text) (Porebski *et al*., 2012). *Mt*DTBS possesses a Gly12, but the equivalent residues (Glu, Asn; underlined text above) from other DTBS are larger and more charged. It is hypothesized that these residues may sterically prevent the promiscuous binding mechanism observed with *Mt*DTBS.

Examples of non-homologous enzymes that exhibited NTP promiscuity identified in the literature differed from *Mt*DTBS, suggesting the mechanism described in chapter 3 is unique. Nucleoside diphosphate kinase (NDK), a phosphotransferase, can bind ADP and TTP (PDB IDs: 1KDN and 1NDC respectively)(Xu *et al*., 1997). Inspection of these crystal structures revealed that NDK binds the aromatic purine and pyrimidine heterocycles through π-π stacking interactions with the side chain of a phenylalanine. Here, the defined interactions between enzyme and nucleoside differentiate NDK promiscuity from that of *Mt*DTBS. The helicase domain from the bacteriophage T7 replicative helicase-primase is able to hydrolyze several NTPs, however, only dTTP and dATP confer helicase activity (Satapathy *et al*., 2009; Sawaya *et al*., 1999). In crystal structures with these ligands, dTTP was modelled into minimal electron density compared to dATP (PDB ID: 1CR1 and 1CR2 respectively). This is similar to the differences between the high affinity and promiscuous binding modes of *Mt*DTBS; CTP was modelled into complete electron density, while only electron density for the triphosphate was observed for the remaining NTPs. Interestingly, the P-loop in the bacteriophage T7 replicative helicase-primase is structurally similar to *Mt*DTBS, and is bound to a sulfate ion in the structure devoid of NTP (PDB ID: 1CR0). Unlike MtDTBS, however, amino acid residues bind to the nucleoside moiety to govern specificity (Satapathy *et al*., 2009). Nucleoside triphosphate diphosphohydrolase isoforms NTPDase1 and NTPDase3 from *Toxoplasma gondii* also differentiate between ATP and ADP substrates through altered binding modes (Krug *et al*., 2013). Structures of NTPDase1 in complex with non-hydrolysable analogues of ATP (AMPPNP; PDB ID: 4KH4) and ADP (AMPNP; PDB ID: 4KH5) revealed some structural data similar to that observed for the promiscuous *Mt*DTBS binding mode, indicated by ligand density solely for the ligand polyphosphate (present in chain A). However, unlike *Mt*DTBS, the alternate subunit (chain B) contained electron density pertaining to the full ligand in both structures, suggesting that this may be due to the other factors, such as the accessibility of chain A due to the crystal packing. Interestingly, NTPDase1 does not bind the triphosphate moiety with a P-loop structure. Finally, protein kinase PKR is activated by RNA containing a 5′-triphosphate moiety, irrespective of the RNA sequence (Toroney *et al*., 2012). This ability appears similar to the promiscuous binding mode for *Mt*DTBS, where NTPs are bound by the 5′-triphosphate, irrespective of the structure of the nucleoside. Therefore, atomic investigation into the

mechanism of RNA triphosphate binding may reveal further comparative insight to the *Mt*DTBS NTP promiscuity mechanism. These select examples detail several similarities to the biophysical capabilities and observed structural data for *Mt*DTBS, including NTP promiscuity, alternate binding modes, the observation of weak nucleoside density in the crystallography data and promiscuity enabled by the binding of a common triphosphate moiety. However, the stark contrast between the alternating binding modes and the nature of crystallographic data observed for the promiscuous binding mode appear to be unique to *Mt*DTBS.

### 5.3.2 Future directions for the investigation *Mt*DTBS promiscuity

Further investigation into the nature of *Mt*DTBS promiscuity would benefit from the use of a technique that observes dynamic phenomena. As mentioned in chapter 4, X-ray crystallographic models are a static snapshot which can differ from what occurs in solution (Davis and Erlanson, 2013). While this technique was crucial to the investigation of *Mt*DTBS promiscuity, there may be further insights that were not observed. X-ray crystallography requires the formation of ordered protein crystals, which is governed by intermolecular interactions between proteins. This is known as crystal packing, and can prevent conformational changes and also result in conformations distinct from those in solution (Davis and Erlanson, 2013). It is possible that the *Mt*DTBS active site is more dynamic in solution and crystal packing restricted this conformational freedom, preventing the observation of specific nucleoside-protein contacts for NTPs binding in the promiscuous binding mode. The SPR data determined that these NTPs exhibited similar affinities despite structural differences, supporting that binding occurred predominantly through the triphosphate tail that is common to all NTPs. However, the affinities of promiscuous NTPs did vary ($K_D$ 16.5 – 93 μM; Chapter 3), suggesting that there were further protein-ligand interactions which contributed to binding but were unobserved. One possible method to test this would be to employ heteronuclear single quantum coherence NMR (HSQC-NMR), as discussed in Chapter 4. Briefly, proteins can be labelled with $^{15}N$, enabling the observation of each nitrogen atom present, including the peptide backbone and certain side chains. Upon ligand binding, the chemical environment of surrounding nitrogens is changed, producing a measurable effect. Therefore, by comparing HSQC-NMR spectra of *Mt*DTBS in the absence and presence of a ligand, it is possible to determine which residues the ligand interacts with. Subsequently, one could compare the difference between spectra of *Mt*DTBS in the presence of CTP and adenosine triphosphate (ATP), for example, to determine if these ligands interact with similar residues. Consequently, interactions additional to those between the triphosphate tail of ATP and the P-loop could be observed, revealing whether there are specific nucleoside-protein interactions that were unable to be observed by crystallography.

An alternate method to further investigate *Mt*DTBS NTP promiscuity would be to identify alternate crystallization conditions that resulted in crystals with different packing. These would most likely require precipitants other than $SO_4^{2-}$ or citrate. It is hypothesized that, upon binding to the P-loop of *Mt*DTBS, these precipitants structurally stabilise this region, which either encourages or allows the formation of the crystal packing interactions. This is based on the observation that *Mt*DTBS crystals rapidly degraded when transferred from a precipitant-containing solution to one devoid of a ligand that could bind at the P-loop, suggesting that loss of precipitant leads to a loss of crystal stability, possibly as a result of abrogated crystal contacts (chapter 2). Although *Mt*DTBS crystals were only observed in crystallization conditions containing anionic precipitants in chapter 2, alternate crystallographic methods such as crystal seeding were not employed. Seeding has been identified as a useful tool for

the optimisation of crystal growth conditions (Bergfors, 2007). This involves the use of crushed, pre-grown crystal fragments to bypass the nucleation stage, allowing crystal growth to occur in alternate conditions. As *Mt*DTBS readily crystallizes in a number of conditions, this would be a promising avenue to identify alternate crystal forms. Ideally, new crystallization conditions would be free from anionic precipitants such as $SO_4^{2-}$, citrate or other organic acids, and result in an unbound P-loop. While a crystal structure of *Mt*DTBS with no ligands modelled at the P-loop has been previously reported (PDB ID: 3FGN), this was not able to be replicated using the same conditions. It is hypothesized that investigation into alternate conditions may reveal novel structural information regarding conformational flexibility of *Mt*DTBS. Furthermore, co-crystallization or soaking with NTPs using an alternate crystal form of *Mt*DTBS may reveal previously unobserved intermolecular interactions, providing further insight into the structural mechanism of NTP promiscuity.

Given that P-loops are commonly observed in ATP and GTP binding proteins throughout the proteome, it is possible that this mechanism of promiscuity will be observed in other proteins. Initial structural and biophysical investigations into NTP promiscuity in other enzymes should be performed on DTBS from other species, in particular that from *Helicobacter pylori*. As described previously, the active site of *Hp*DTBS bears similarities to that of *Mt*DTBS; a highly positively charged area at the P-loop and a lack of hydrogen bonding to the ribose ring of its bound NTP. Investigation of *Hp*DTBS NTP binding would determine if these specific structural features contribute to the unique mechanism of *Mt*DTBS NTP promiscuity. Subsequently, studying other P-loop containing proteins from throughout the proteome may provide further insight into the occurrence of this phenomenon and the structural features that cause it. Examples of ligand promiscuity in the literature may be difficult to find due to the use of different identifying terms; other researchers may have discovered similar phenomena but not referred to it as promiscuity. Therefore, a more thorough method of identifying other potential proteins that bind NTPs promiscuously would be to search protein databases (e.g. NCBI, RCSB PDB, Phyre2) for similar sequences or structures to the *Mt*DTBS P-loop. After identifying several proteins, their genes could be transformed into bacteria for expression and purification, followed by biophysical and structural investigation as performed in chapters 2 and 3. To conclude, investigation into NTP utilization in DTBS from other species will aid the determination of the structural features required for promiscuity. Comparison of the activities of enzymes that bind ligands promiscuously may also provide insight into the biological purpose of this phenomenon. As was hypothesized for *Mt*DTBS, this may be a method to retain enzyme activity at all times, regardless of nutrient availability. A practical use of identifying the minimal structural requirements of NTP promiscuity would be to engineer other enzymes to utilize a range of energy sources, which may be beneficial for industrial purposes where it is important to utilize cheaper reagents.

### 5.3.3 *In vivo* implications of NTP promiscuity

Despite biochemical, biophysical and structural characterization of NTP promiscuity, the effect of this phenomenon within the mycobacterium remains unclear. In chapter 3, it was determined that the *Mt*DTBS exhibited significantly higher affinity for CTP than other NTPs. Subsequently in chapter 4, it was determined that 2′-deoxycytidine triphosphate (dCTP) also binds to *Mt*DTBS in the same high affinity binding mode exhibited by CTP, but with approximately 3-fold higher affinity. This encourages the use of the available *Mt*DTBS enzyme activity assay to determine the extent to which dCTP is able to be used as an energy source during catalysis. However, despite the high affinity preference for CTP, previous

enzyme kinetic data revealed similar $K_m$ values (CTP $K_m$ = 25.2 ± 1.7 μM, other NTPs $K_m$ = 17 – 30 μM) and a slight $k_{cat}$ preference for CTP (CTP $k_{cat}$ = 15.2 ± 0.3 μM, other NTPs $K_{cat}$ = 10.2 – 11.8 μM) (Salaemae *et al*., 2015). The similarity in these values can be explained by the relatively high affinity of CDP, which could slow the binding of new substrates after each round of catalysis, resulting in an overall rate comparable to the other, lower affinity NTPs. However, it appears that there is little catalytic difference between the high affinity and promiscuous NTPs. Consequently, it is unclear as to what degree NTP promiscuity affects energy source use by *Mt*DTBS during catalysis within the mycobacterium. We have previously hypothesized that NTP promiscuity allows *Mt*DTBS to remain active at all stages of the *Mtb* lifecycle, owing to the fact that the gene is essential for survival (Salaemae *et al*., 2015) (Chapter 3). However, given the vast affinity differences between NTP binding modes (~2,000-fold) and an appreciable amount of CTP or dCTP, *Mt*DTBS would be most likely bound to these substrates rather than any of the promiscuous NTPs. Indeed, it may be that the levels of cytosine-containing nucleoside triphosphates available to *Mt*DTBS depreciates significantly throughout the lifecycle, causing it to rely on other substrates. Although it is known that the concentrations of NTP vary throughout the lifecycle of bacteria (Buckstein *et al*., 2008; Conlon *et al*., 2016), this has not been studied in *Mtb*. In order to fully understand the biological reasoning behind the evolution of NTP promiscuity, investigation of NTP levels at varying stages of the complex *Mtb* lifestyle is required. This could be performed via quantitative measurement of NTP levels from *Mtb* whole cell lysates at differing stages of the lifecycle, and under certain stressors such as antibiotic treatment. Using this information in combination with the affinities determined in this thesis, the NTP most likely to be bound to *Mt*DTBS at different stages of the lifecycle can be mathematically modelled. While it is possible that *Mt*DTBS is required to use alternate sources of energy, it may be that a high affinity substrate such as dCTP is present a sufficient level to be ubiquitously utilized for catalysis. Such studies may reveal insight into whether or not *Mt*DTBS has evolved NTP promiscuity as a biological requirement to maintain activity.

### 5.3.4 The adenosine diphosphate binding mode

Throughout this thesis, there have been several observations that invite further investigation into the intermolecular interactions between *Mt*DTBS and adenosine diphosphate (ADP). In chapter 2 it was observed that ADP bound in a conformation distinct from that of CTP, with the adenosine moiety exhibiting two binding poses, one within and one adjacent to the DAPA pocket. Electron density distinguished the entirety of the ADP in this crystal structure, which was also distinct from that of adenosine triphosphate (ATP) in chapter 3, where electron density was only present for the triphosphate moiety. This comparison raised the possibility that the adenosine moiety of ATP may also be able to bind in the poses observed for ADP. However, it was modelled in a state that would allow for catalytic activity, an ability that was previously determined by enzyme activity assays (Salaemae *et al*., 2015). In chapter 3, it was also observed that ADP exhibited a higher affinity for *Mt*DTBS than ATP (17.9 μM vs 75 μM respectively). This was unusual, considering that the addition of a phosphate caused a 3-fold increase in affinity between CDP and CTP. However, the structural data is consistent with the difference in affinity between ADP and ATP; electron density for the whole of ADP was present, indicating that it forms a more stable complex. Additionally, the alternate ADP binding modes exhibited new interactions that could have caused an increase in affinity. Adjacent to the DAPA pocket, extra hydrogen bonds were observed between the adenosine moiety and *Mt*DTBS. Within the predominantly hydrophobic DAPA pocket, it is likely that several hydrophobic forces contributed to binding the adenosine moiety. In both modes, hydrogen bonds between the ribose ring and *Mt*DTBS were also observed. As mentioned

previously, these interactions could be investigated and validated by alternate structural techniques, such as HSQC-NMR. Specifically, comparison of spectra performed in the presence of ADP or ATP would reveal whether or not these ligands interact with the same *Mt*DTBS residues. Likewise, comparison between ADP and DAPA would reveal whether the adenosine moiety interacts with residues within the DAPA pocket. Confirming these interactions by an orthogonal method would bring additional validity to the binding mode observed by crystallography.

Altogether, the presented data regarding the ADP binding mode are consistent, inviting speculation into the biological relevance of these findings. Naturally, after a certain number of catalytic cycles where ATP is used as the energy source, the concentration of ADP would increase. The higher affinity of ADP suggests that relatively few catalytic cycles would have to occur before ADP begins to outcompete ATP. However, *Mt*DTBS is able to catalyze its reaction using ATP, with similar $K_m$ and $k_{cat}$ values observed for other NTPs (all NTPs $K_m$ = *17 – 30* μM; $K_{cat}$ = 10.2 – 15.2 $s^{-1} \times 10^{-3}$) (Salaemae *et al*., 2015). Although it is known that ADP is an inhibitor of *Mt*DTBS, these $K_m$ and $k_{cat}$ values demonstrate that ADP does not cause product inhibition to the point where catalysis is prematurely prevented. A possible explanation for this is that when ADP binds in small amounts as a result of catalytic turnover, it is outcompeted by the simultaneous binding of both ATP and DAPA. Indeed, phosphoryl transfer between ATP and DAPA has been shown in *Ec*DTBS to include additional residues, which would likely increase the affinity of the overall complex (Huang *et al*., 1995; Käck *et al*., 1998). Further investigation is required to determine the biological implications of the observed ADP binding modes; however, it is clear that ATP can be competently utilized by *Mt*DTBS for catalysis.

## 5.4 Fragment-based lead discovery and design against *Mt*DTBS

### 5.4.1 *Mt*DTBS is an amenable drug discovery target

*Mt*DTBS is a promising target for the development of anti-tuberculosis drugs. Foremost, because the biotin biosynthesis pathway is essential to *Mtb* infection and survival (Dey *et al*., 2010; Salaemae *et al*., 2011, 2016; Woong Park *et al*., 2011). Additional to this, there now exists an increasing body of knowledge regarding DTBS from multiple species, including *Mt*DTBS (Salaemae *et al*., 2015), *Ec*DTBS (Huang *et al*., 1995; Käck *et al*., 1998) and *Hp*DTBS (Porebski *et al*., 2012). Comparison to DTBS from other species is not only useful for the investigation of natural ligand binding, it will also aid the development of molecules that bind specifically to *Mt*DTBS and only inhibit *Mtb*. Indeed, focusing on the development of narrow-spectrum antibiotics has been suggested to be the future of infectious disease medicine, especially in regards to combating resistance (Brown and Wright, 2016). In addition to this, there are no human homologues to *Mt*DTBS, indicating that off-target effects will be minimal. Finally, and importantly, *Mt*DTBS is highly amenable to biochemical, biophysical and structural investigation. A consideration that can be overlooked, this has enabled the development of several robust protocols which constitute a pipeline for the accurate and rapid assaying of compounds or inhibitors. *Mt*DTBS is readily immobilized to a streptavidin coated SPR chip, retaining activity for up to 3 weeks (chapter 3). Using this assay, it was possible to rapidly, accurately and precisely investigate the protein-ligand interactions with over 20 compounds, including nucleoside polyphosphates, fragments and analogues. It is clear that further SPR experiments, such as fragment screening, are possible with this enzyme. *Mt*DTBS also meets all of the ‘fundamental crystal requirements’ for

protein amenability to crystallography in fragment-based projects (Öster *et al*., 2015). These requirements were: robustness, including reproducible crystal growth, high stability, adequate diffraction; an accessible binding site; and 'soakability'. *Mt*DTBS readily forms crystals in a range of conditions that contain cryo-protectants, taking roughly 3 days to begin growing and reaching a full size of 100 – 400 µM and consistently diffracting to sufficient resolution for atomic analysis. *Mt*DTBS crystals are highly stable, and are able to withstand high concentration soaking experiments and physical manipulation, such as during precipitant-ligand exchange. Generally, the maximum concentration of DMSO tolerated by *Mt*DTBS crystals was 10% (v/v), corresponding to high concentrations of ligands. The ligand-bound structure of compound **2** in chapter 4 required a concentration of at least 30 mM; it is possible that this interaction would not have been observed had high concentrations not been possible. Additionally, NTPs were present at concentrations of ~300 mM during the structural investigations into promiscuity. The use of extremely high fragment concentrations during soaking (20 – 500 mM) has been demonstrated to be highly useful in FBLD projects (Bauman *et al*., 2016). Indeed, there is also evidence suggesting that high concentration crystallographic soaking results in the highest hit-rates, identifying many hits that other primary screening techniques would not (Schiebel *et al*., 2016). The amenability of *Mt*DTBS to biophysical and structural investigation, especially high concentration crystallographic soaking, further exemplifies this enzyme as a promising target for drug discovery projects. Foremost, this highlights the feasibility of performing fragment screening by SPR and X-ray crystallography. These experiments were determined to be lower priority than the characterisation of *Mt*DTBS natural ligand binding (Chapters 2, 3) and the identification and optimisation of the hits (Chapter 4) presented in this thesis, but should be higher priority for the continuation of this project.

### 5.4.2 *Mt*DTBS fragment screening by crystallography

As discussed in chapter 4, further fragments need to be identified in order aid future drug discovery against *Mt*DTBS. While some successful development was performed with the two cytidine analogues and one fragment that were identified to bind to *Mt*DTBS, further probing of the active site would increase the capacity for growing, merging and linking of existing compounds. The Booker laboratory has access to a ChemBridge fragment library consisting of 1,135 compounds, which is rule of three compliant (MW ≤300, H-bond donors ≤3, H-bond acceptors ≤3, cLogP ≤3)(Carr *et al*., 2005). This library is stored at Compounds Australia. In order to aid the use of this library for fragment screening by crystallography, it would be preferable to have compounds solubilized to as high a concentration as possible. This is because, as discussed previously, high concentrations are preferable for fragment soaks, and the amount of compound that can be tolerated by crystals is generally limited by the concentration of DMSO (~10% for *Mt*DTBS crystals). Considering this, there is a 'High Solubility Set' of fragments within the ChemBridge library that are soluble in 100% DMSO to at least 200 mM, which could result in soaking concentrations of at least 20 mM. Using the Art Robbins Phoenix crystallography drop setting robot, it would be possible to input up to 288 independent sitting drop crystallization conditions per tray (Intelliplate). Compounds can then be soaked into each well that contains a crystal and incubated for a period of time. Given that the reported *Mt*DTBS crystal conditions contain cryo-protectants, each crystal can be mounted and flash cooled in liquid nitrogen with minimal handling. Subsequently, the crystals can be subjected to X-ray diffraction at the Australian Synchrotron. Datasets are automatically autoprocessed using XDS (Kabsch, 2010) at the synchrotron, and each one should be subject to the following criteria before continuing: <3 Å, $CC_{1/2}$ > 0.3 and completeness > 0.85. Following this, truncation and phasing can be automated using CCP4

(Winn *et al.*, 2011) and in-house *Mt*DTBS search models. Finally, PanDDA automatically identifies areas of electron density that may pertain to bound fragments (Pearce *et al.*, 2017). By comparing multiple datasets, PanDDA is also able to enhance the electron density of weakly bound fragments or those that are obscured by water molecules. Following this, comparison to fragment screens using orthogonal techniques such as SPR should be performed. Overall, fragment screening by X-ray crystallography is feasible with *Mt*DTBS as the target protein.

There are several considerations for the use of *Mt*DTBS in fragment screening. Firstly, the P-loop is bound to the precipitant molecules $SO_4^{2-}$ or citrate in *Mt*DTBS crystal structures. Given that these precipitants prevented the binding of NTPs, it is likely that they would also prevent the binding of fragments to this region. One option would be to perform precipitant-ligand exchange with fragments, however this would require a large amount of potentially error prone manual crystal handling. Furthermore, fragments may not bind tightly enough to prevent crystal degradation. Another option would be to identify alternate crystallization conditions for *Mt*DTBS that do not contain anionic precipitants. Nevertheless, despite the presence of $SO_4^{2-}$ or citrate, fragment screening would still be beneficial to discover new compounds that bind to the cytidine binding site, DAPA binding pocket or even other, hereto unknown *Mt*DTBS binding sites. The next consideration is one of library design. X-ray crystallographic screening requires highly soluble hits, which can be determined computationally with cLogP, a value that has shown to have a weak correlation to hit success rates (Öster *et al.*, 2015). By further filtering the fragment library based on a value such as this, the number of compounds required to screen would reduce. This would reduce the number of individual crystals required to mount, which is likely a major bottle-neck to throughput of screening by X-ray crystallography and one of the predominant reasons why this was not attempted in the current project. Indeed, X-ray fragment screening at Diamond Light Source uses robotic assistance for crystal harvesting (Oxford University, UK, 2018). Without such equipment, compounds can be soaked in pools, rather than in singleton. With careful consideration of the factors presented here, a fragment screen of *Mt*DTBS using X-ray crystallography would be a useful and feasible exercise to identify alternate compounds that may aid future development of inhibitors.

### 5.4.3 Advances in the characterisation of the *Mt*DTBS active site

Throughout this thesis, the *Mt*DTBS active site was probed by biophysical and structural methods, revealing key characteristics that must be taken into account when designing inhibitors. These will be discussed subsequently, starting with the cytidine binding site and followed by the P-loop and the DAPA pocket.

Cytidine binds into a shallow, solvent exposed pocket on the surface of *Mt*DTBS. The cytosine base in particular binds extremely efficiently to this pocket, forming 4 hydrogen bonds with the peptide backbone of residues in the region. Throughout this thesis, no alternate moieties were observed to bind into cytosine pocket, despite testing several alternate NTPs and cytidine analogues via crystallography. While this observation demonstrates the specificity of the interaction between cytosine and *Mt*DTBS, it is still possible that alternate molecules are able to bind here. This is especially so when considering the possibility that the *Mt*DTBS active site may differ in solution from what is observed by crystallography. Unlike the cytosine pocket, it appears that the ribose moiety of CTP does not exhibit hydrogen bonds to the enzyme. However, structural and biophysical evidence indicates that the 2′-hydroxyl group of CTP interacts unfavourably with nearby hydrophobic residues, leading to slower

association and faster dissociation than the unsubstituted dCTP. This highlights that future development of compounds intended to bind in the CTP pocket may benefit from SAR of the ribose moiety, in order to determine the optimal substituents. It is possible that the entire ribose moiety could be replaced by a more efficient chemical structure. Finally, the cytidine binding pocket is likely crucial to the development of narrow-spectrum anti-tuberculosis drugs which, as mentioned previously, is an important consideration for future antibiotic development (Brown and Wright, 2016). As observed in chapter 3, this region is structurally distinct between DTBS from different species. Furthermore, it appears that the preference for CTP is unique to *Mt*DTBS. Therefore, targeting the cytidine binding pocket may achieve specificity over DTBS from other species.

Located between the cytidine binding site and the DAPA pocket, the P-loop is a short motif that binds to the triphosphates of NTPs with high affinity. The LRLR nest clearly has a high propensity for binding anionic species, and was identified as crucial to the mechanism by which *Mt*DTBS elicits NTP promiscuity. No developed compounds were observed to bind to the LRLR nest in this thesis, but this must be a primary goal for the future of compound design against *Mt*DTBS. Whether as an analogue from the **2**-series or with a linked compound, performing chemical growth of a compound to interact with the P-loop would provide a significant boost to affinity.

The DAPA pocket is a large hydrophobic cavity that forms at the homo-dimer interface between *Mt*DTBS monomers. The natural substrate DAPA binds to this pocket with a moderate affinity ($K_D$ = 80 μM ± 20 μM; Chapter 3), which suggests that outcompeting this substrate with a high affinity compound is feasible. However, DAPA is a relatively simple alkyl chain compound, with a small molecular weight and heavy atom count, resulting in quite an efficient molecule (MW = 188 Da; Ligand efficiency = 0.43 vs 0.32 for compound **5**). Thus, DAPA itself may constitute a preferable moiety to include in future inhibitor design, particularly within a linked compound. Nonetheless, it appears that the DAPA pocket is able to accommodate a variety of different moieties, such as the **2**-series compounds and ADP. These observations invite further probing of this region, particularly with a fragment screen.

### 5.4.4 Future directions for the development of *Mt*DTBS inhibitors

There are currently two compound series that have been identified and developed: the **2**-series and the linked compound series. Each series will be discussed in the context of the development of inhibitors of the enzyme and the mycobacterium. Firstly, the location at which the compound binds within the active site will determine which substrate it competes with *in vivo*. Indeed, it appears that **2**-series compounds bind to the same region of the *Mt*DTBS active site as the ɣ-phosphate of CTP, indicating that the two would compete for binding. This was one hypothesis as to why **2**-series compounds did not inhibit *Mtb*, and it was suggested that higher affinity for *Mt*DTBS is required. Intended inhibitors should exhibit higher affinity than any competing natural substrates. Conversely, it may be beneficial to build **2**-series compounds away from where CTP binds, such that they only compete with DAPA. However, the low affinity of **2**-series compounds compared to CTP supports the development of a linked compound. Designing a molecule that binds optimally to the cytidine binding site, P-loop and DAPA pocket would likely exhibit extremely high affinity. However, care must be taken to ensure that linked compounds exhibit the appropriate drug-like characteristics. Indeed, many anti-TB drugs are smaller and more hydrophilic molecules, expected to enter the mycobacterium through porin channels (Lambert, 2002). Rifampicin, on

the other hand, is more akin to the large linked compounds (**9** MW = 524 Da; **10** MW = 628 Da; rifampicin MW = 823 Da). However, the current linked compounds are less lipophilic than rifampicin, which is suggested to be one of the reasons it can passively traverse the mycobacterial cell wall (Lambert, 2002) (**9** LogP = 0.51, **10** LogP = -2.32, rifampicin LogP = 2.62; values determined using Molinspiration). Future development of linked compounds to become more lipophilic may aid cell penetration in the future. Nevertheless, the iterative compound development process will require continual testing against the mycobacterium, which is performed by collaborators at the Prof. James Triccas laboratory (University of Sydney, Australia). Indeed, bacterial cell penetration of drugs has been identified as one of the key challenges for the development of new antibiotics (Lamoree and Hubbard, 2018; Payne *et al*., 2007; Tommasi *et al*., 2015). In order to address this, there is on-going research into the mechanisms behind cell permeability and to monitor the accumulation of antibiotic compounds within bacteria (Masi *et al*., 2017). Along with structural, biophysical and biochemical assays, including techniques that observe the bacterial cell permeability of compounds during the iterative development of *Mt*DTBS inhibitors would also greatly increase the chances of successful inhibition of *Mtb*.

## 5.5 Concluding remarks

The work presented in this thesis represents the progression of a project aiming to develop inhibitors for the promising anti-tuberculosis drug target *Mt*DTBS. By using a fragment-based approach, accompanied by careful consideration and investigation into the biophysical nature of the enzyme, a significant improvement in the binding characteristics of novel compounds was achieved. This highlights that FBLD is an enabling method, encouraging participation in drug discovery by researchers who were previously prohibited by infrastructural barriers. It is hypothesized that the iterative FBLD optimization process will suit the development of future antibiotic compounds. Using the techniques reported here, it will be possible to rapidly, accurately and precisely guide compound development using biochemical, biophysical and structural information. Regular inclusion of assays to test cell permeability and bacterial inhibition will likewise guide the development of new anti-tubercular compounds. Such innovations in drug discovery will provide a strong and positive influence on the successful treatment of tuberculosis and the rise in antibiotic resistance.

# Appendices

**Appendix 1; Table 1. Data collection and refinement statistics.**

| | dCTN | dCDP | dCTP | dFdC | 2 | Synthesized-2 |
|---|---|---|---|---|---|---|
| **Wavelength (Å)** | 0.9537 | 0.7108 | 0.9537 | 0.9537 | 0.9537 | 0.9537 |
| **Resolution range** | 49.01 - 2.23 (2.31 - 2.23) | 46.18 - 2.0 (2.071 - 2.0) | 45.83 - 2.125 (2.201 - 2.125) | 38.35 - 1.991 (2.062 - 1.991) | 46.29 - 1.9 (1.968 - 1.9) | 38.55 - 2.4 (2.486 - 2.4) |
| **Space group** | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ |
| **Unit cell (a b c (Å) α β ɣ (°))** | 55.2 106.2 152.9 90 90 90 | 54.8 105.2 154.2 90 90 90 | 54.4 104.1 153.1 90 90 90 | 55.5 106.1 152.0 90 90 90 | 54.6 105.7 154.5 90 90 90 | 56.1 106.2 152.8 90 90 90 |
| **Total reflections** | 89236 (8645) | 651931 (62556) | 30546 (2990) | 123750 (11711) | 1904288 (189745) | 707995 (66092) |
| **Unique reflections** | 44653 (4102) | 57086 (4923) | 25126 (2574) | 62008 (5910) | 71394 (6988) | 36160 (3140) |
| **Multiplicity** | 2.0 (2.0) | 11.4 (10.7) | 1.2 (1.2) | 2.0 (2.0) | 26.7 (26.9) | 19.6 (19.4) |
| **Completeness (%)** | 98.68 (87.57) | 87.66 (81.06) | 50.40 (52.50) | 99.34 (96.51) | 99.35 (96.96) | 95.83 (88.08) |
| **Mean I/sigma(I)** | 13.99 (3.62) | 31.20 (4.05) | 8.36 (1.69) | 14.28 (2.14) | 76.12 (3.92) | 28.93 (2.09) |
| **Wilson B-factor** | 25.54 | 22.18 | 42.84 | 32.58 | 22.73 | 48.3 |
| **R-merge** | 0.02848 (0.3063) | 0.7795 (1.953) | 0.01738 (0.1616) | 0.02526 (0.2679) | 0.756 (2.351) | 0.8552 (2.964) |
| **R-meas** | 0.04028 (0.4332) | 0.8154 (2.055) | 0.02458 (0.2286) | 0.03573 (0.3788) | 0.7709 (2.396) | 0.8758 (3.039) |
| **R-pim** | 0.02848 (0.3063) | 0.235 (0.6278) | 0.01738 (0.1616) | 0.02526 (0.2679) | 0.1491 (0.4608) | 0.1852 (0.6556) |
| **CC1/2** | 0.997 (0.772) | 0.799 (0.708) | 0.999 (0.938) | 0.999 (0.913) | 0.838 (0.845) | 0.868 (0.543) |
| **CC*** | 0.999 (0.933) | 0.942 (0.91) | 1 (0.984) | 1 (0.977) | 0.955 (0.957) | 0.964 (0.839) |
| **Reflections used in refinement** | 44138 (3875) | 53601 (4879) | 25125 (2575) | 61885 (5910) | 70939 (6837) | 34984 (3140) |
| **Reflections used for R-free** | 2185 (194) | 2706 (251) | 1218 (118) | 3090 (305) | 3490 (330) | 1711 (155) |
| **R-work** | 0.2012 (0.6390) | 0.2039 (0.3351) | 0.1762 (0.2905) | 0.1979 (0.2844) | 0.2285 (0.5854) | 0.2054 (0.2900) |
| **R-free** | 0.2486 (0.6796) | 0.2716 (0.4490) | 0.2373 (0.3722) | 0.2505 (0.3270) | 0.2761 (0.5861) | 0.2855 (0.3640) |
| **CC(work)** | 0.963 (0.117) | 0.840 (0.601) | 0.963 (0.574) | 0.948 (0.591) | 0.834 (0.413) | 0.747 (0.553) |
| **CC(free)** | 0.955 (0.001) | 0.798 (0.271) | 0.919 (0.302) | 0.924 (0.464) | 0.827 (0.407) | 0.742 (0.461) |
| **Number of non-hydrogen atoms** | 6938 | 7345 | 6724 | 7269 | 7207 | 7066 |
| **macromolecules** | 6326 | 6293 | 6351 | 6382 | 6370 | 6719 |
| **ligands** | 84 | 99 | 112 | 74 | 220 | 164 |
| **solvent** | 528 | 953 | 261 | 813 | 617 | 183 |
| **Protein residues** | 908 | 907 | 908 | 908 | 909 | 908 |
| **RMS(bonds)** | 0.006 | 0.007 | 0.007 | 0.007 | 0.08 | 0.042 |
| **RMS(angles)** | 0.83 | 0.89 | 0.85 | 0.88 | 0.85 | 0.99 |
| **Ramachandran favored (%)** | 97.67 | 98.33 | 98.33 | 98.55 | 98.33 | 97.44 |
| **Ramachandran allowed (%)** | 1.78 | 1 | 1.33 | 1.11 | 1.34 | 1.78 |
| **Ramachandran outliers (%)** | 0.56 | 0.67 | 0.33 | 0.33 | 0.33 | 0.78 |
| **Rotamer outliers (%)** | 2.64 | 3.15 | 2.95 | 1.79 | 1.14 | 3.51 |
| **Clashscore** | 3.8 | 3.48 | 5.82 | 3.53 | 6.23 | 8.75 |
| **Average B-factor** | 37.03 | 24.39 | 45.85 | 37.02 | 39.47 | 60.3 |
| **macromolecules** | 36.81 | 23.8 | 45.92 | 36.29 | 39.43 | 60.68 |
| **ligands** | 43.13 | 23.44 | 40.62 | 47.03 | 36.39 | 58.85 |
| **solvent** | 38.62 | 28.32 | 46.55 | 41.87 | 40.97 | 47.67 |

Statistics for the highest-resolution shell are shown in parentheses. Values are subject to change prior to PDB submission.
Tables were made using Table 1 (PHENIX), with the most recently refined .pdb and .mtz files, as well as the relevant unmerged .mtz

**Appendix 1; Table 2. Data collection and refinement statistics continued.**

| | 2 & CTN | 3a | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| **Wavelength (Å)** | 1.54 | 0.9537 | 0.9537 | 0.9537 | 0.9537 | 0.9537 |
| **Resolution range** | 53.28 - 2.6 (2.693 - 2.6) | 45.97 - 1.94 (2.009 - 1.939) | 37.69 - 2.3 (2.382 - 2.3) | 46.34 - 2.438 (2.525 - 2.438) | 44.46 - 2.652 (2.747 - 2.652) | 45.77 - 1.691 (1.751 - 1.691) |
| **Space group** | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ |
| **Unit cell (a b c (Å) α β ɣ (°))** | 54.5 106.6 154.3 90 90 90 | 56.3 106.9 152.8 90 90 90 | 55.6 108.4 153.8 90 90 90 | 54.8 105.6 152.5 90 90 90 | 55.1 105.2 150.5 90 90 90 | 53.9 105.0 154.3 90 90 90 |
| **Total reflections** | 875992 (82802) | 137720 (13152) | 1414302 (127424) | 66739 (5798) | 52194 (5062) | 195343 (18390) |
| **Unique reflections** | 28191 (2632) | 68914 (6618) | 41902 (3779) | 33416 (2935) | 26106 (2538) | 98148 (9363) |
| **Multiplicity** | 31.1 (31.0) | 2.0 (2.0) | 33.8 (32.3) | 2.0 (2.0) | 2.0 (2.0) | 2.0 (2.0) |
| **Completeness (%)** | 98.10 (95.78) | 99.57 (97.05) | 90.45 (85.88) | 98.53 (88.16) | 99.84 (98.83) | 99.44 (96.29) |
| **Mean I/sigma(I)** | 59.82 (3.53) | 18.31 (2.81) | 76.96 (2.20) | 21.16 (2.88) | 11.71 (2.25) | 11.79 (1.66) |
| **Wilson B-factor** | 51.91 | 35.87 | 39.46 | 52.4 | 43.93 | 21.07 |
| **R-merge** | 0.7296 (1.811) | 0.01317 (0.173) | 0.6221 (9.689) | 0.0163 (0.2101) | 0.03916 (0.2362) | 0.03214 (0.3269) |
| **R-meas** | 0.7411 (1.841) | 0.01862 (0.2446) | 0.6313 (9.848) | 0.02305 (0.2972) | 0.05538 (0.3341) | 0.04546 (0.4623) |
| **R-pim** | 0.1281 (0.3225) | 0.01317 (0.173) | 0.1056 (1.704) | 0.0163 (0.2101) | 0.03916 (0.2362) | 0.03214 (0.3269) |
| **CC1/2** | 0.904 (0.677) | 1 (0.953) | 0.982 (0.709) | 1 (0.951) | 0.998 (0.854) | 0.999 (0.839) |
| **CC*** | 0.974 (0.898) | 1 (0.988) | 0.995 (0.911) | 1 (0.987) | 1 (0.96) | 1 (0.955) |
| **Reflections used in refinement** | 27878 (2631) | 68899 (6618) | 38125 (3546) | 33364 (2935) | 26105 (2538) | 97991 (9363) |
| **Reflections used for R-free** | 1385 (110) | 3448 (309) | 1893 (159) | 1736 (137) | 1377 (138) | 4978 (463) |
| **R-work** | 0.2010 (0.2771) | 0.1991 (0.2676) | 0.2456 (0.3336) | 0.1987 (0.2768) | 0.1839 (0.2173) | 0.1973 (0.2598) |
| **R-free** | 0.2880 (0.3502) | 0.2388 (0.3234) | 0.3015 (0.3636) | 0.2629 (0.3610) | 0.2574 (0.3380) | 0.2298 (0.2955) |
| **CC(work)** | 0.848 (0.527) | 0.947 (0.651) | 0.766 (0.687) | 0.933 (0.572) | 0.952 (0.788) | 0.956 (0.693) |
| **CC(free)** | 0.862 (0.614) | 0.929 (0.619) | 0.743 (0.678) | 0.917 (0.326) | 0.936 (0.691) | 0.941 (0.643) |
| **Number of non-hydrogen atoms** | 6618 | 7027 | 6912 | 6793 | 6848 | 7775 |
| **macromolecules** | 6295 | 6364 | 6560 | 6549 | 6580 | 6580 |
| **ligands** | 141 | 118 | 108 | 100 | 56 | 96 |
| **solvent** | 182 | 545 | 244 | 144 | 212 | 1099 |
| **Protein residues** | 907 | 908 | 909 | 909 | 909 | 909 |
| **RMS(bonds)** | 0.043 | 0.007 | 0.045 | 0.033 | 0.007 | 0.039 |
| **RMS(angles)** | 0.98 | 0.85 | 1.02 | 0.91 | 0.9 | 0.86 |
| **Ramachandran favored (%)** | 95.55 | 98.56 | 98.44 | 97.11 | 97.55 | 98.44 |
| **Ramachandran allowed (%)** | 3.23 | 1.22 | 1.22 | 2.44 | 2 | 1.22 |
| **Ramachandran outliers (%)** | 1.22 | 0.22 | 0.33 | 0.44 | 0.45 | 0.33 |
| **Rotamer outliers (%)** | 5.46 | 1.96 | 3.14 | 2.83 | 2.97 | 1.72 |
| **Clashscore** | 8.73 | 6.04 | 8.27 | 6.65 | 6.77 | 4.75 |
| **Average B-factor** | 56.63 | 42.16 | 48.75 | 57.1 | 44.1 | 25.6 |
| **macromolecules** | 56.94 | 41.81 | 49.08 | 57.41 | 44.23 | 24.42 |
| **ligands** | 53.21 | 49.63 | 46.12 | 46.81 | 42.66 | 16.85 |
| **solvent** | 48.46 | 44.67 | 40.99 | 49.99 | 40.25 | 33.44 |

Statistics for the highest-resolution shell are shown in parentheses. Values are subject to change prior to PDB submission.
Tables were made using Table 1 (PHENIX), with the most recently refined .pdb and .mtz files, as well as the relevant unmerged .mtz

**Appendix 1; Table 3. Data collection and refinement statistics continued.**

| | 8 | 9 | 10 |
|---|---|---|---|
| **Wavelength (Å)** | 0.9537 | 0.9537 | 0.9537 |
| **Resolution range** | 46.13 - 1.99 (2.061 - 1.99) | 45.51 - 2.44 (2.528 - 2.44) | 45.99 - 2.005 (2.077 - 2.005) |
| **Space group** | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ | $P\ 2_12_12_1$ |
| **Unit cell (a b c (Å) α β ɣ (°))** | 54.4 104.3 154.3 90 90 90 | 56.8 106.0 151.2 90 90 90 | 54.7 104.6 153.6 90 90 90 |
| **Total reflections** | 122013 (11854) | 69313 (6673) | 119699 (11578) |
| **Unique reflections** | 61101 (5975) | 34674 (3354) | 59871 (5806) |
| **Multiplicity** | 2.0 (2.0) | 2.0 (2.0) | 2.0 (2.0) |
| **Completeness (%)** | 99.57 (99.02) | 99.80 (98.47) | 99.75 (98.01) |
| **Mean I/sigma(I)** | 13.57 (2.00) | 17.45 (2.85) | 17.55 (2.72) |
| **Wilson B-factor** | 29.03 | 41.47 | 36.41 |
| **R-merge** | 0.03028 (0.2679) | 0.02354 (0.1977) | 0.01501 (0.1871) |
| **R-meas** | 0.04282 (0.3789) | 0.03329 (0.2796) | 0.02123 (0.2646) |
| **R-pim** | 0.03028 (0.2679) | 0.02354 (0.1977) | 0.01501 (0.1871) |
| **CC1/2** | 0.999 (0.889) | 0.999 (0.947) | 1 (0.928) |
| **CC*** | 1 (0.97) | 1 (0.986) | 1 (0.981) |
| **Reflections used in refinement** | 60902 (5975) | 34668 (3354) | 59869 (5806) |
| **Reflections used for R-free** | 3106 (304) | 1737 (163) | 3026 (312) |
| **R-work** | 0.2108 (0.2742) | 0.1883 (0.2602) | 0.1922 (0.2777) |
| **R-free** | 0.2499 (0.3222) | 0.2510 (0.3409) | 0.2277 (0.3286) |
| **CC(work)** | 0.943 (0.676) | 0.930 (0.578) | 0.954 (0.681) |
| **CC(free)** | 0.936 (0.582) | 0.895 (0.493) | 0.936 (0.628) |
| **Number of non-hydrogen atoms** | 7255 | 6711 | 6907 |
| **macromolecules** | 6580 | 6263 | 6351 |
| **ligands** | 40 | 69 | 58 |
| **solvent** | 635 | 379 | 498 |
| **Protein residues** | 909 | 905 | 908 |
| **RMS(bonds)** | 0.016 | 0.007 | 0.007 |
| **RMS(angles)** | 1.81 | 0.88 | 0.85 |
| **Ramachandran favored (%)** | 98.22 | 98.22 | 98.44 |
| **Ramachandran allowed (%)** | 1.34 | 1.45 | 1.22 |
| **Ramachandran outliers (%)** | 0.45 | 0.33 | 0.33 |
| **Rotamer outliers (%)** | 1.56 | 4.49 | 2.3 |
| **Clashscore** | 2.84 | 6.64 | 3.63 |
| **Average B-factor** | 32.33 | 41.81 | 40.12 |
| **macromolecules** | 31.64 | 41.85 | 39.91 |
| **ligands** | 27.87 | 46.92 | 34.8 |
| **solvent** | 39.77 | 40.08 | 43.49 |

Statistics for the highest-resolution shell are shown in parentheses. Values are subject to change prior to PDB submission.
Tables were made using Table 1 (PHENIX), with the most recently refined .pdb and .mtz files, as well as the relevant unmerged .mtz

## Appendix 2: Surface plasmon resonance data

Contained in this appendix is the remaining SPR ligand binding data referred to in Chapter 4. Included are the kinetic parameters determined for dCDP and dCTP (Table 1). See Chapter 4, Figure 4.4 for example sensorgrams with these ligands. Subsequently, the steady state analysis parameters and example sensorgrams for compounds **2**-**10** are included (Table 2, Figure 1).

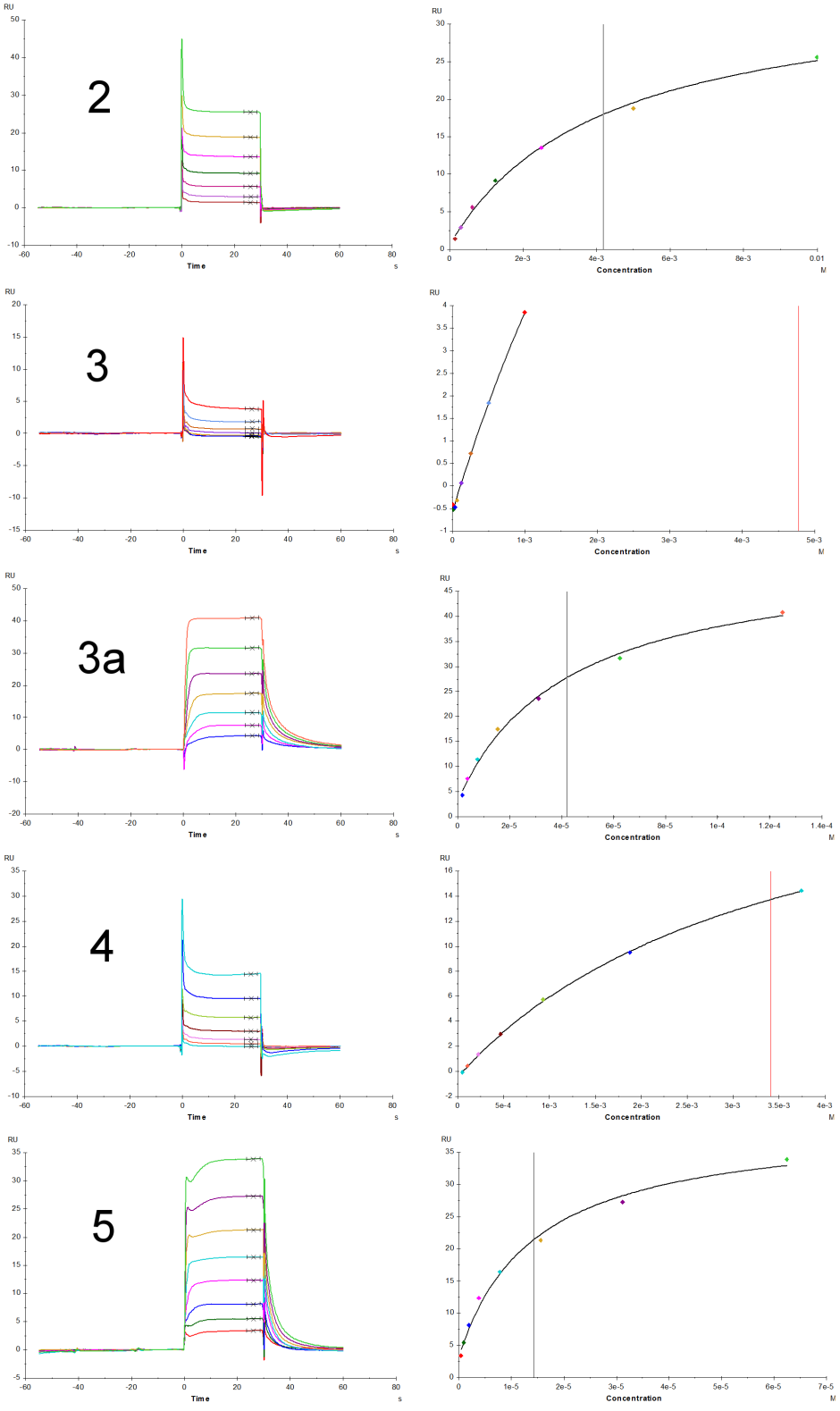
**Table 1:** Kinetics parameters for dCDP and dCTP SPR data. Data is collated from three experiments.

| Sample | $k_a$ (1/Ms) | SEM ($k_a$) (1/Ms) | $k_d$ (1/s) | SEM ($k_d$) (1/s) | $K_D$ (M) | Rmax (RU) | Chi² (RU²) | Replicates |
|---|---|---|---|---|---|---|---|---|
| dCDP | 1.2e+6 | 4e+5 | 0.12 | 0.03 | 1.18e-7 | 30-50 | 0.02 – 0.1 | 3 |
| dCTP | 7.13e+5 | 0.05e+5 | 0.0341 | 0.0002 | 4.76e-8 | 53 | 0.5 | 3 |

SEM: standard error of the mean

**Table 2:** Steady state affinity analysis and averaged $K_D$ with standard error of the mean for compounds **2**-**10**.

| Substrate | $K_D$ (μM) | Rmax (RU) | Chi² (RU²) |
|---|---|---|---|
| **2** | 2700 | 48 | 1.05 |
| | 2700 | 39 | 0.01 |
| | 4100 | 39 | 0.52 |
| | 4200 | 35 | 0.05 |
| Average | 3400 ± 400 | | |
| **3** | 2000 | 21 | 0.01 |
| | 5000 | 26 | 0.01 |
| Average | 4000 ± 1000 | | |
| **3a** | 50 | 52 | 0.45 |
| | 47 | 50 | 0.93 |
| | 42 | 50 | 1.14 |
| Average | 46 ± 2 | | |
| **4** | 1000 | 16 | 0.14 |
| | 6000 | 47 | 0.19 |
| | 3000 | 28 | 0.01 |
| Average | 4000 ± 1000 | | |
| **5** | 28 | 40 | 3.56 |
| | 14 | 36 | 1.19 |
| | 14 | 37 | 1.19 |
| Average | 19 ± 5 | | |
| **6** | 16 | 32 | 0.39 |
| | 19 | 34 | 1.16 |
| | 13 | 33 | 0.22 |
| | 19 | 36 | 0.90 |
| Average | 17 ± 2 | | |
| **7** | 42 | 38 | 0.96 |
| | 32 | 39 | 5.95 |
| | 40 | 39 | 0.55 |
| | 40 | 35 | 2.51 |
| Average | 38 ± 2 | | |
| **8** | 72 | 43 | 0.45 |
| | 50 | 36 | 0.22 |
| | 47 | 49 | 1.89 |
| Average | 56 ± 8 | | |
| **9** | 400 | 36 | 0.09 |
| | 800 | 49 | 0.01 |
| Average | 600 ± 200 | | |
| **10** | 800 | 39 | 0.29 |
| | 800 | 64 | 0.55 |
| | 300 | 53 | 0.61 |
| Average | 700 ± 200 | | |


2
RU
Time
s
RU
Concentration
M
3
RU
Time
s
RU
Concentration
M
3a
RU
Time
s
RU
Concentration
M
4
RU
Time
s
RU
Concentration
M
5
RU
Time
s
RU
Concentration
M

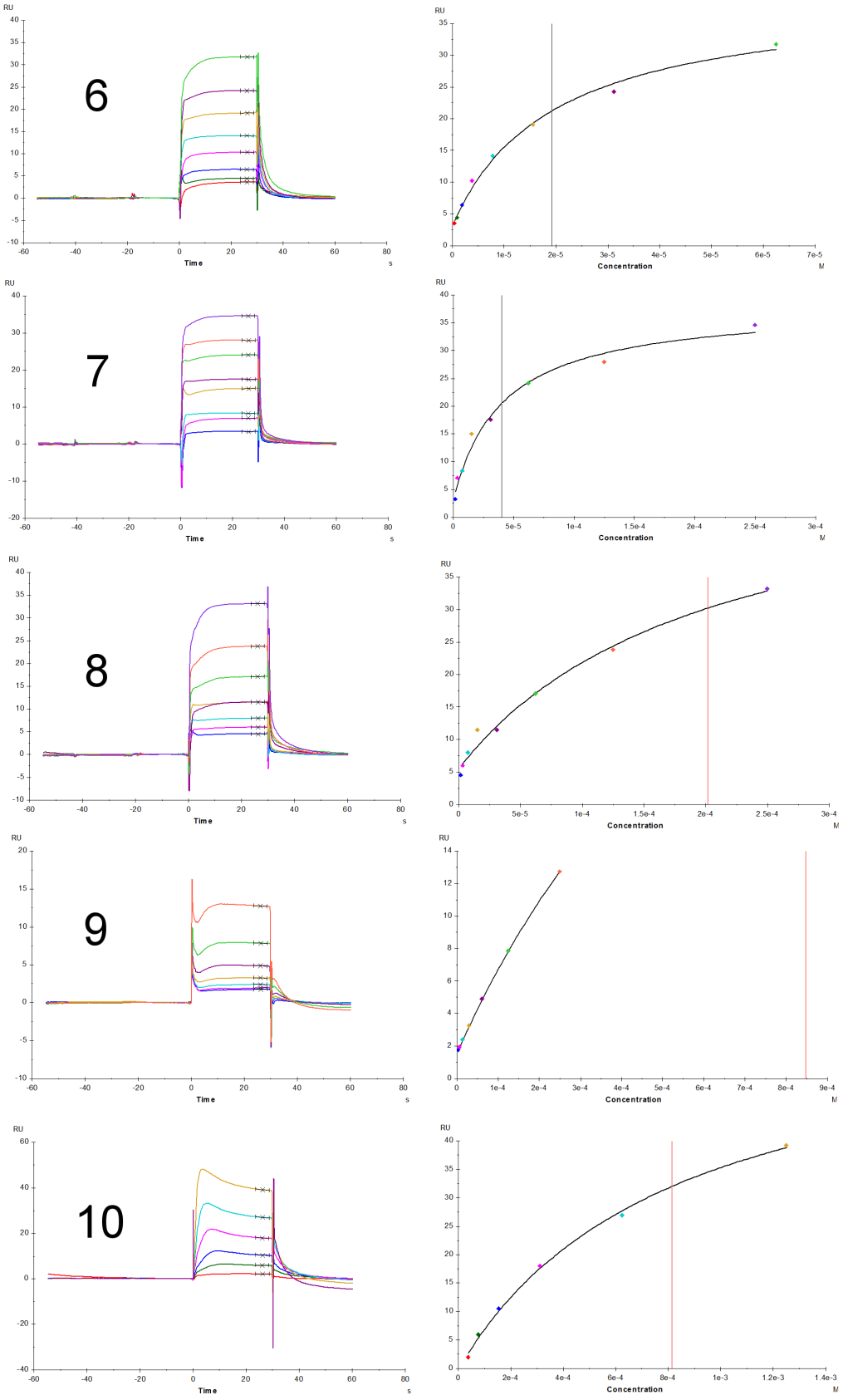


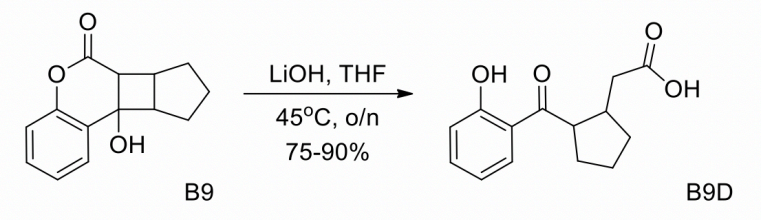

**Figure 1:** Examples of SPR sensorgrams for ligands **2-10** (left). Fast kinetics were observed and the ligands were fit with the steady-state affinity model (Right; Biacore S200 Evaluation Software).

Enclosed within this appendix is select data that is relevant to the current thesis. This work was adapted from the *Mt*DTBS Linking Project Report by Birgit Gaiser. Please note: Compounds **1** and **2** are referred to as per the notebook designations B9 and B9D respectively throughout this report.

**B9 degradation**

*1) B9 under basic conditions*

B9 was treated with LiOH in THF/$H_2O$ at 45°C and the formation of B9D was confirmed by NMR and MS (Scheme 1).




**Scheme 1**. Induced degradation of B9 to B9D.

*2) B9 under neutral aqueous conditions*

A sample of B9 in DMSO had been stored on the bench for several years. Presumably, the sample had become wet and the water has induced degradation. Over time the solvent had evaporated and the residue contained a crystal and an amorphous solid. The crystal was analysed by small molecule X-ray crystallography and was shown to be B9 (vida infra). The amorphous residue was analysed by NMR spectroscopy, showing that the sample consisted in large parts of B9, but that B9D was also present. The NMR spectrum indicated a ratio of ca. 2:1 in favour of B9, but taken into consideration the B9 crystal that had been removed before the NMR analysis, the ratio was approximately 3:1.

*3) B9 under crystallization conditions*

As it was confirmed in the above described experiment that B9 was still the major component of the sample for binding assays and crystallization experiments, the question arose what had happened to B9 in the crystallization soak. Therefore, a sample of the crystallization buffer (1.4 M $NH_4SO_4$, 0.1 M Tris pH 8, 10% glycerol) containing B9 was analysed by TLC and it was shown that B9D formation had started after 2 days and no B9 was left after 6 days. However, the sample had not been checked on days 4 and 5.

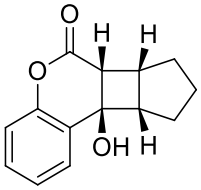
These experiments show that B9 only slowly degrades to B9D under neutral aqueous conditions, but that the degradation gets faster under increasing basic conditions.

**Characterization of B9 and B9D / Stereochemistry**

*B9*

B9 was known to be racemic, but the supplier could not specify the stereochemistry of the compound. B9 was characterized in-house in order to shine light on its stereochemistry. NMR, HPLC and MS confirmed the structure of B9 and optical rotation confirmed that it is a racemic mixture ($[\alpha]_D^{25}$ +0.800).

NOE experiments could unfortunately not unravel the stereochemistry. However, the crystal mentioned above was analysed by small molecule X ray crystallography. B9 was shown to be the cis – cis – cis isomer (Figure 1).



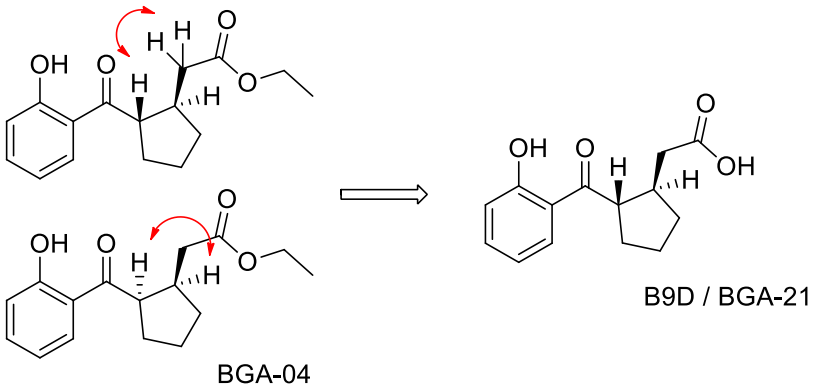

*rac*-B9

**Figure 1.** B9 structure confirmed by small molecule X-ray crystallography.

*B9D*

B9D (Scheme 1) was analysed in the same means as B9, however not by X-ray crystallography. As for B9, the structure was confirmed, but NOE experiments could not unravel the stereochemistry of B9D.

The relative stereochemistry of B9D and its analogues was finally determined with the aid of compound BGA-04 (Figure 2). This ethyl ester was synthesized by mistake (vida infra) and surprisingly consisted of two separable diastereoisomers. Both diastereoisomers were analysed by NMR. By comparison of their NOE spectra, peak 1 was identified as the trans isomer and peak 2 as the cis isomer. The trans isomer showed a correlation between the α-proton and the $CH_2$ group, whereas the cis isomer showed a correlation between the two single protons on the cyclopentane ring (Figure 2). Notably, the α-proton undergoes a rather large downfield shift when switching from trans to cis ($\delta_{cis} \approx 3.9$ ppm; $\delta_{trans} \approx 3.5$ ppm). Comparison of the $^1H$ NMR spectra of the two BGA-04 peaks with the one of B9D led to the conclusion that B9D is in its trans conformation.




BGA-04 B9D / BGA-21

**Figure 2.** Structure of the two BGA-04 diastereoisomers (left top: trans-BGA-04; left bottom: cis-BGA04) and structure of B9D (right). Red arrows indicate the correlations found in NOESY experiments.

Having obtained this knowledge, the NMR spectrum of the B9/DMSO sample was compared to the NMR of BGA-21 and by this, it was revealed that the B9D contained in this sample was mostly *trans*-B9D (ca. 75 %), but *cis*-B9D was also present (ca. 25 %).

In summary, the stereochemistry of B9D could not be determined until the B9D ethyl ester (BGA-04) was synthesized by mistake and luckily was obtained as a mixture of separable diastereoisomers. The fact that the B9/DMSO sample contained mostly *trans*-B9D indicates that the *cis*-isomer is poorly stable, even under neutral conditions. However, BGA-04 was synthesized under basic conditions, and yet the *cis*-isomer was obtained.

## Appendix 4: Adelaide Proteomics Centre LC/MS compound analysis

**Data acquisition – Orbitrap – Liquid Chromatography ESI-MS (LC-ESI-MS).**

The compound was diluted 1:5 with water prior to LC-ESI-MS analysis. Nano-LC−ESI-MS/MS was performed on an Ultimate 3000 RSLC system (Thermo-Fisher Scientific) coupled to a LTQ Orbitrap XL ETD MS instrument (both Thermo-Fisher Scientific). 20 µL of the sample was pre-concentrated onto a C18 trapping column (Acclaim PepMap100 C18 75 µm × 20 mm, Thermo-Fisher Scientific) at a flow rate of 5 µL/min in 2% ACN 0.1% TFA for 5 minutes. Separation was performed using a 75 µm ID C18 column (Acclaim PepMap100 C18 75 µm × 15 cm, Thermo-Fisher Scientific) at a flow rate of 0.3 µL/min using a linear gradient from 5 to 90% B (A: 5% ACN 0.1% FA, B: 80% ACN 0.1% FA) over 30 minutes, followed by a 10 minute wash with 95% B, and an 20 minute equilibration with 5%. MS scans were acquired in the mass range of 50 *m/z* to 500 *m/z* at a resolution of 60,000.

**Data analysis**

Data analysis was performed using the XCalibur software (Version 2.0.7, Thermo Fisher Scientific). The acquired spectrum was compared with a theoretical spectrum using the elemental composition of the compounds utilizing the isotope simulation option of the XCalibur software.

Report prepared by Georgia Arentz and Yin Ying Ho (Adelaide Proteomics Centre).

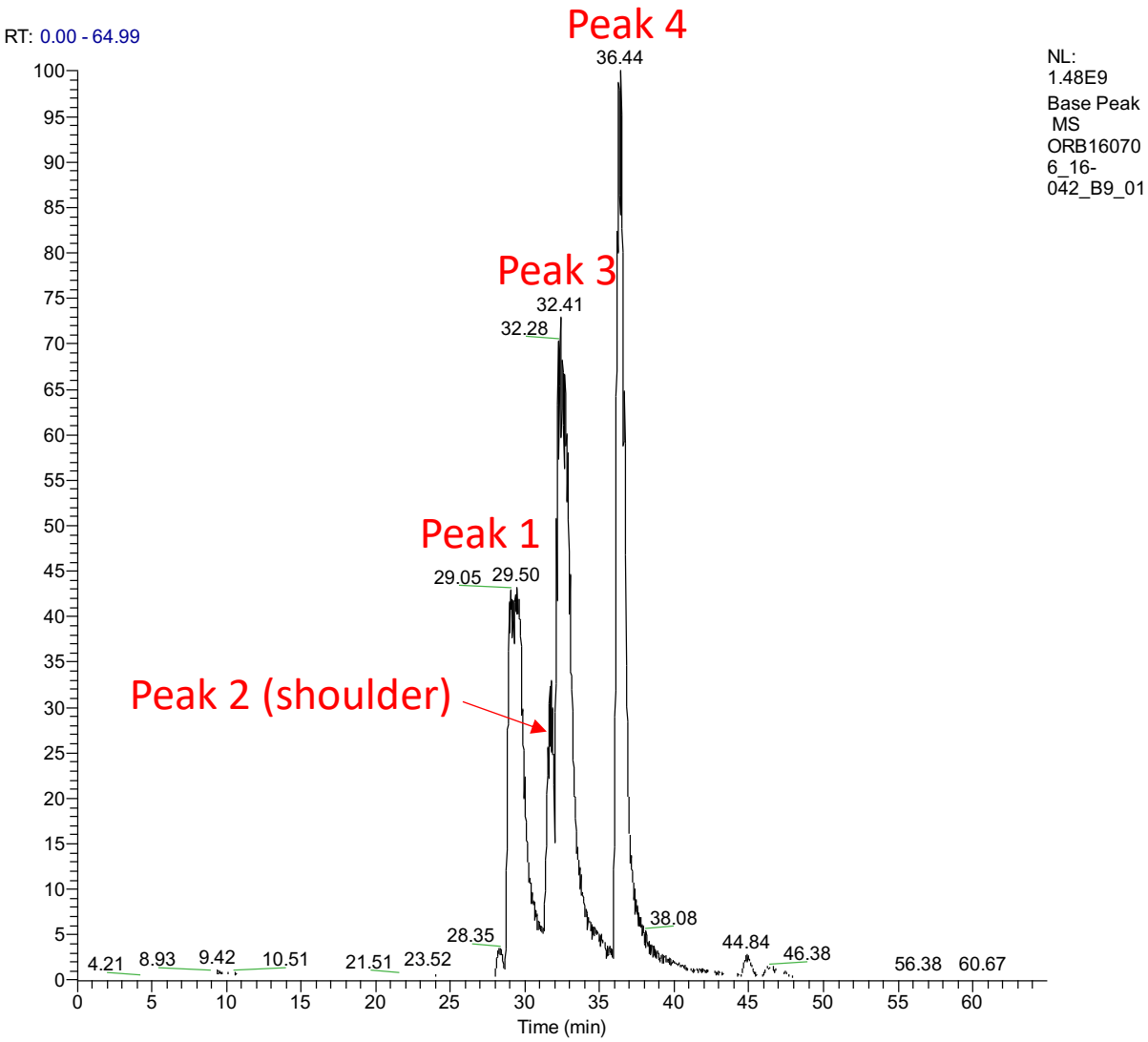



**Figure 1:** Total ion chromatography for the stock sample of compound **1**. Four independent peaks are observed.

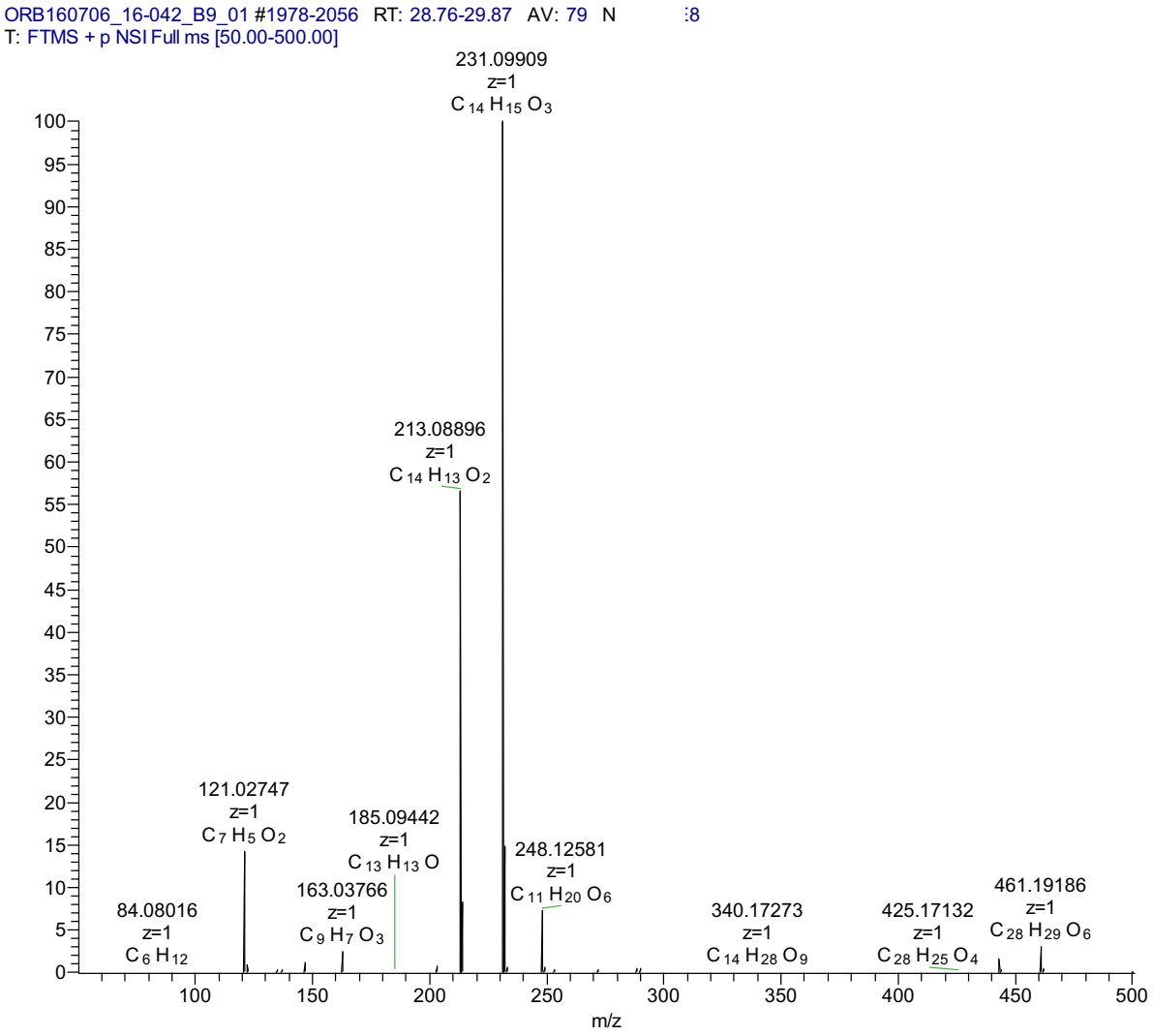



**Figure 2**: MS spectrum for analysis of peak 1. *M/Z* pertaining to **1** (231), **2** (248) and an unidentified compound (213) were observed.

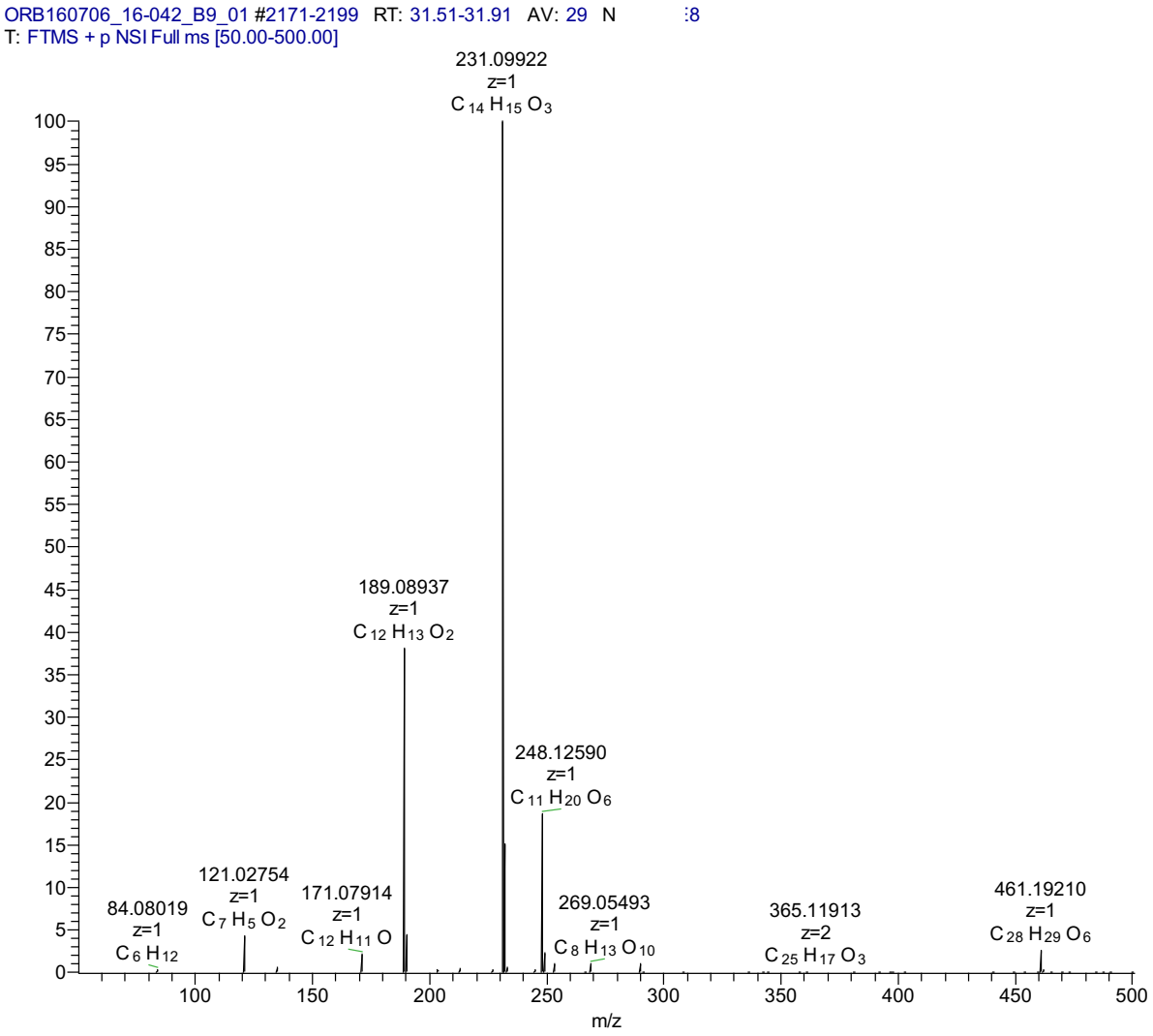



**Figure 3:** MS spectrum for analysis of peak 2 (shoulder). *M/Z* pertaining to **1** (231), **2** (248), as well as an unidentified compound (189) were observed.

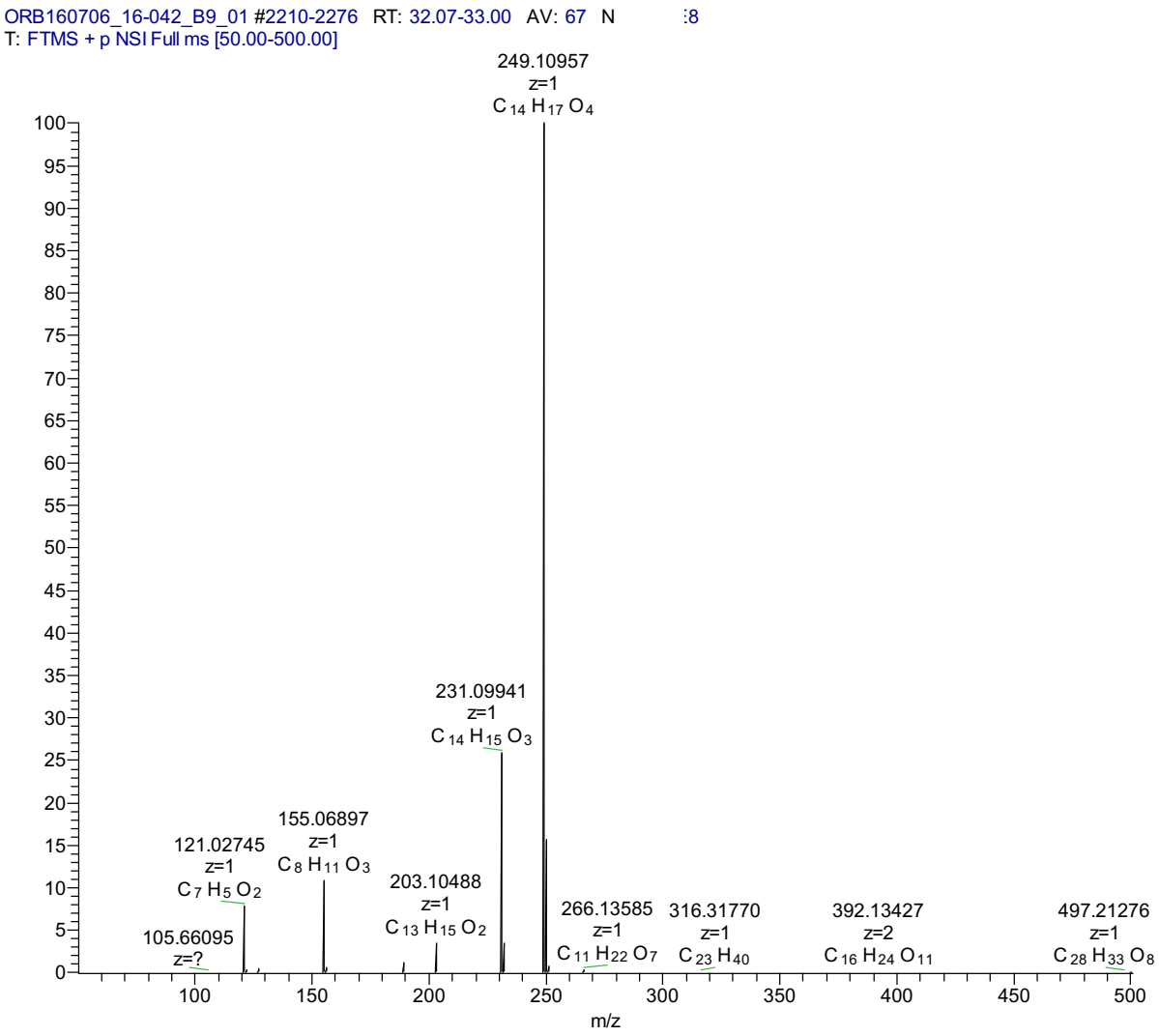



**Figure 4:** MS spectrum for analysis of peak 3. *M/Z* pertaining to **1** (231) and **2** (249) were observed. The variation in the *m/z* of **2** was attributed to differing ionization states.

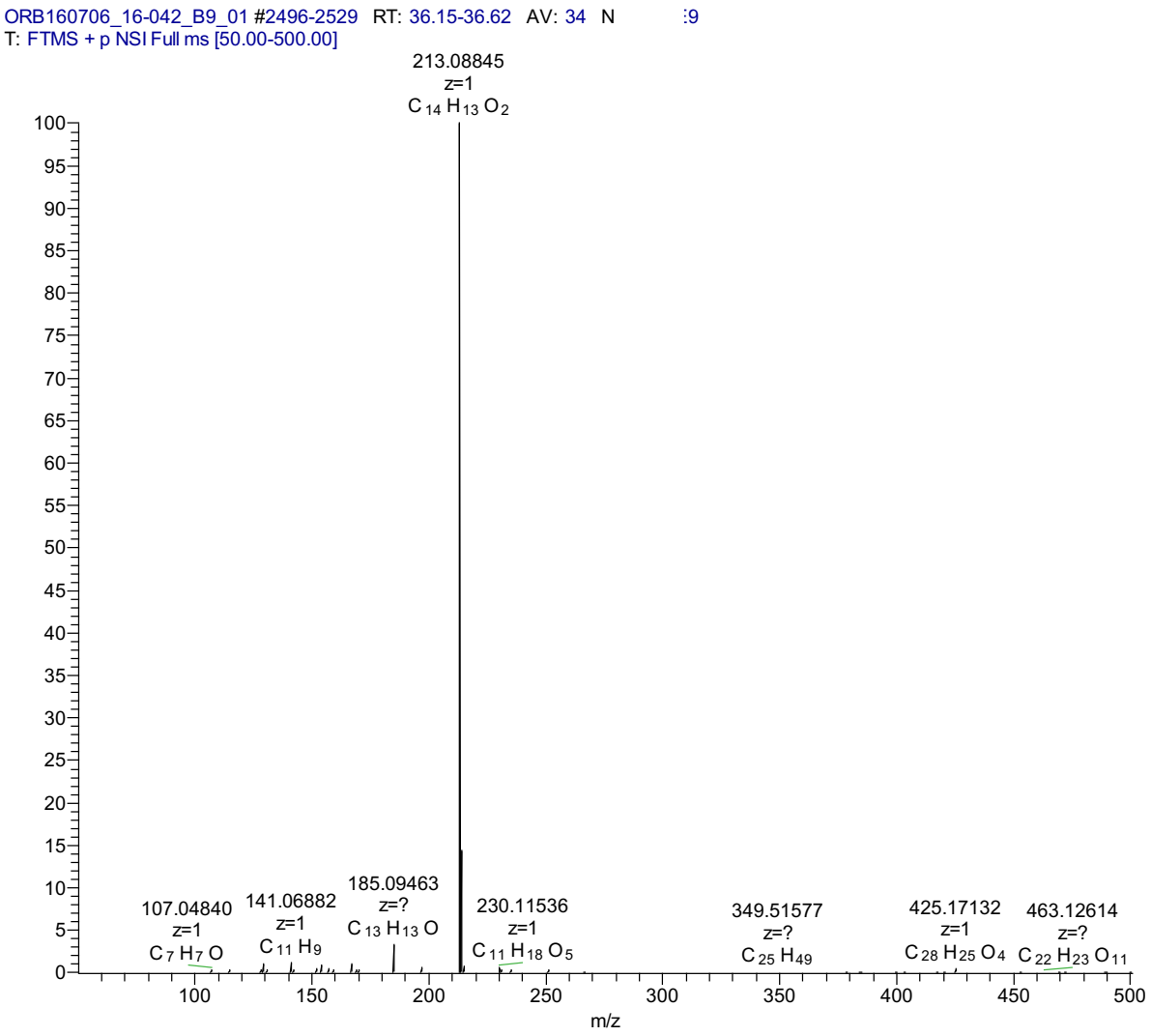



**Figure 5:** MS spectrum for analysis of peak 4. *M/Z* pertaining to an unidentified compound (213) was observed. This compound corresponds to **1** with a loss of water.

# References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst D *66*, 213–221.

Alvarellos, M.L., Lamba, J., Sangkuhl, K., Thorn, C.F., Wang, L., Klein, D.J., Altman, R.B., and Klein, T.E. (2014). PharmGKB summary: Gemcitabine Pathway. Pharmacogenet Genomics *24*, 564–574.

Andrews, P.R., Craik, D.J., and Martin, J.L. (1984). Functional group contributions to drug-receptor interactions. J. Med. Chem. *27*, 1648–1657.

Baker, M. (2013). Fragment-based lead discovery grows up. Nat. Rev. Drug Discov. *12*, 5–7.

Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Cryst D *67*, 271–281.

Bauman, J.D., Harrison, J.J.E.K., and Arnold, E. (2016). Rapid experimental SAD phasing and hot-spot identification with halogenated fragments. IUCrJ *3*, 51–60.

Bergfors, T. (2007). Succeeding with seeding: some practical advice. In Evolving Methods for Macromolecular Crystallography, R.J. Read, and J.L. Sussman, eds. (Springer Netherlands), pp. 1–10.

Brown, E.D., and Wright, G.D. (2016). Antibacterial drug discovery in the resistance era. Nature *529*, 336–343.

Buckstein, M.H., He, J., and Rubin, H. (2008). Characterization of Nucleotide Pools as a Function of Physiological State in Escherichia coli. J. Bacteriol. *190*, 718–726.

Carr, R.A.E., Congreve, M., Murray, C.W., and Rees, D.C. (2005). Fragment-based lead discovery: leads by design. Drug Discov. Today *10*, 987–992.

Chang, C.A., Chen, W., and Gilson, M.K. (2007). Ligand configurational entropy and protein binding. PNAS *104*, 1534–1539.

Chung, S., Parker, J.B., Bianchet, M., Amzel, L.M., and Stivers, J.T. (2009). Impact of linker strain and flexibility in the design of a fragment-based inhibitor. Nat Chem Biol *5*, 407–413.

Coimbra, J.R.M., Marques, D.F.F., Baptista, S.J., Pereira, C.M.F., Moreira, P.I., Dinis, T.C.P., Santos, A.E., and Salvador, J.A.R. (2018). Highlights in BACE1 Inhibitors for Alzheimer's Disease Treatment. Front Chem *6*.

Conlon, B.P., Rowe, S.E., Gandt, A.B., Nuxoll, A.S., Donegan, N.P., Zalis, E.A., Clair, G., Adkins, J.N., Cheung, A.L., and Lewis, K. (2016). Persister formation in *Staphylococcus aureus* is associated with ATP depletion. Nat. Microbiol. *1*, 16051.

Davis, B.J., and Erlanson, D.A. (2013). Learning from our mistakes: The 'unknown knowns' in fragment screening. Bioorg. Med. Chem. Lett. *23*, 2844–2852.

Dey, S., Lane, J.M., Lee, R.E., Rubin, E.J., and Sacchettini, J.C. (2010). Structural Characterization of the Mycobacterium tuberculosis Biotin Biosynthesis Enzymes 7,8-Diaminopelargonic Acid Synthase and Dethiobiotin Synthetase. Biochemistry *49*, 6746–6760.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Cryst D *60*, 2126–2132.

Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W., and Jhoti, H. (2016). Twenty years on: the impact of fragments on drug discovery. Nat. Rev. Drug Discov. *15*, 605–619.

Feng, J., Paparella, A.S., Tieu, W., Heim, D., Clark, S., Hayes, A., Booker, G.W., Polyak, S.W., and Abell, A.D. (2016). New Series of BPL Inhibitors To Probe the Ribose-Binding Pocket of Staphylococcus aureus Biotin Protein Ligase. ACS Med. Chem. Lett. *7*, 1068–1072.

Greenhill, J.V., and Lue, P. (1993). 5 Amidines and Guanidines in Medicinal Chemistry. In Progress in Medicinal Chemistry, G.P. Ellis, and D.K. Luscombe, eds. (Elsevier), pp. 203–326.

Hopkins, A.L., Keserü, G.M., Leeson, P.D., Rees, D.C., and Reynolds, C.H. (2014). The role of ligand efficiency metrics in drug discovery. Nat. Rev. Drug Discov. *13*, 105–121.

Huang, W., Jia, J., Gibson, K.J., Taylor, W.S., Rendina, A.R., Schneider, G., and Lindqvist, Y. (1995). Mechanism of an ATP-dependent carboxylase, dethiobiotin synthetase, based on crystallographic studies of complexes with substrates and a reaction intermediate. Biochemistry *34*, 10985–10995.

Ichihara, O., Barker, J., Law, R.J., and Whittaker, M. (2011). Compound Design by Fragment-Linking. Mol. Inform. *30*, 298–306.

Jencks, W.P. (1981). On the attribution and additivity of binding energies. Proc. Natl. Acad. Sci. U.S.A. *78*, 4046–4050.

Joseph-McCarthy, D., Campbell, A.J., Kern, G., and Moustakas, D. (2014). Fragment-based lead discovery and design. J. Chem. Inf. and Model. *54*, 693–704.

Kabsch, W. (2010). XDS. Acta Cryst D *66*, 125–132.

Käck, H., Gibson, K.J., Lindqvist, Y., and Schneider, G. (1998). Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. PNAS *95*, 5495–5500.

Karplus, P.A., and Diederichs, K. (2012). Linking Crystallographic Model and Data Quality. Science *336*, 1030–1033.

Karplus, P.A., and Diederichs, K. (2015). Assessing and maximizing data quality in macromolecular crystallography. Curr Opin Struct Biol *34*, 60–68.

Krug, U., Totzauer, R., Zebisch, M., and Sträter, N. (2013). The ATP/ADP Substrate Specificity Switch between Toxoplasma gondii NTPDase1 and NTPDase3 is Caused by an Altered Mode of Binding of the Substrate Base. ChemBioChem *14*, 2292–2300.

Lambert, P.A. (2002). Cellular impermeability and uptake of biocides and antibiotics in Gram-positive bacteria and mycobacteria. J. Appl. Microbiol. *92*, 46S-54S.

Lamoree, B., and Hubbard, R.E. (2018). Using Fragment-Based Approaches to Discover New Antibiotics. SLAS Discov *23*, 495–510.

Lockhart, S.R., Fothergill, A.W., Iqbal, N., Bolden, C.B., Grossman, N.T., Garvey, E.P., Brand, S.R., Hoekstra, W.J., Schotzinger, R.J., Ottinger, E., et al. (2016). The Investigational Fungal Cyp51 Inhibitor VT-1129 Demonstrates Potent In Vitro Activity against Cryptococcus neoformans and Cryptococcus gattii. Antimicrob. Agents and Chemother. *60*, 2528–2531.

Mashalidis, E.H., Śledź, P., Lang, S., and Abell, C. (2013). A three-stage biophysical screening cascade for fragment-based drug discovery. Nat. Protoc. *8*, 2309–2324.

Masi, M., Réfregiers, M., Pos, K.M., and Pagès, J.-M. (2017). Mechanisms of envelope permeability and antibiotic influx and efflux in Gram-negative bacteria. Nat. Microbiol. *2*, 17001.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658–674.

McPhillips, T.M., McPhillips, S.E., Chiu, H.-J., Cohen, A.E., Deacon, A.M., Ellis, P.J., Garman, E., Gonzalez, A., Sauter, N.K., Phizackerley, R.P., et al. (2002). Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. J Synchrotron Radiat *9*, 401–406.

Myznikov, L.V., Hrabalek, A., and Koldobskii, G.I. (2007). Drugs in the tetrazole series. (Review). Chem Heterocycl Com *43*, 1–9.

Öster, L., Tapani, S., Xue, Y., and Käck, H. (2015). Successful generation of structural information for fragment-based drug discovery. Drug Discov. Today *20*, 1104–1111.

Paparella, A.S., Feng, J., Blanco-Rodriguez, B., Feng, Z., Phetsang, W., Blaskovich, M.A.T., Cooper, M.A., Booker, G.W., Polyak, S.W., and Abell, A.D. (2018a). A template guided approach to generating cell permeable inhibitors of Staphylococcus aureus biotin protein ligase. Tetrahedron *74*, 1175–1183.

Paparella, A.S., Lee, K.J., Hayes, A.J., Feng, J., Feng, Z., Cini, D., Deshmukh, S., Booker, G.W., Wilce, M.C.J., Polyak, S.W., et al. (2018b). Halogenation of Biotin Protein Ligase Inhibitors Improves Whole Cell Activity against Staphylococcus aureus. ACS Infect. Dis. *4*, 175–184.

Payne, D.J., Gwynn, M.N., Holmes, D.J., and Pompliano, D.L. (2007). Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nat. Rev. Drug Discov. *6*, 29–40.

Pearce, N.M., Krojer, T., Bradley, A.R., Collins, P., Nowak, R.P., Talon, R., Marsden, B.D., Kelm, S., Shi, J., Deane, C.M., et al. (2017). A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. Nat. Comm. *8*, 15123.

Porebski, P.J., Klimecka, M., Chruszcz, M., Nicholls, R.A., Murzyn, K., Cuff, M.E., Xu, X., Cymborowski, M., Murshudov, G.N., Savchenko, A., et al. (2012). Structural characterization of H. pylori dethiobiotin synthetase reveals differences between family members. FEBS J *279*, 1093–1105.

Rendina, A.R., Taylor, W.S., Gibson, K., Lorimer, G., Rayner, D., Lockett, B., Kranis, K., Wexler, B., Marcovici-Mizrahi, D., Nudelman, A., et al. (1999). The design and synthesis of inhibitors of dethiobiotin synthetase as potential herbicides. Pesticide Science *55*, 236–247.

Salaemae, W. (2015). Novel inhibitors for biotin biosynthesis pathway in Mycobacterium tuberculosis. Thesis. University of Adelaide.

Salaemae, W., Azhar, A., Booker, G.W., and Polyak, S.W. (2011). Biotin biosynthesis in Mycobacterium tuberculosis: physiology, biochemistry and molecular intervention. Protein Cell *2*, 691–695.

Salaemae, W., Yap, M.Y., Wegener, K.L., Booker, G.W., Wilce, M.C.J., and Polyak, S.W. (2015). Nucleotide triphosphate promiscuity in Mycobacterium tuberculosis dethiobiotin synthetase. Tuberculosis *95*, 259–266.

Salaemae, W., Booker, G.W., and Polyak, S.W. (2016). The Role of Biotin in Bacterial Physiology and Virulence: a Novel Antibiotic Target for Mycobacterium tuberculosis. Microbiol Spectr *4*.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: a laboratory manual (Cold Spring Harbor Laboratory Press).

Satapathy, A.K., Crampton, D.J., Beauchamp, B.B., and Richardson, C.C. (2009). Promiscuous Usage of Nucleotides by the DNA Helicase of Bacteriophage T7. J Biol Chem *284*, 14286–14295.

Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C., and Ellenberger, T. (1999). Crystal Structure of the Helicase Domain from the Replicative Helicase-Primase of Bacteriophage T7. Cell *99*, 167–177.

Schiebel, J., Radeva, N., Krimmer, S.G., Wang, X., Stieler, M., Ehrmann, F.R., Fu, K., Metz, A., Huschmann, F.U., Weiss, M.S., et al. (2016). Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. ACS Chem. Biol. *11*, 1693–1701.

Schramm, V.L. (2018). Enzymatic Transition States and Drug Design. Chem. Rev. *118*, 11194–11258.

Shah, P., and Westwell, A.D. (2007). The role of fluorine in medicinal chemistry. J Enzyme Inhib. Med. Chem. *22*, 527–540.

Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. Science *274*, 1531–1534.

Soares da Costa, T.P., Tieu, W., Yap, M.Y., Pendini, N.R., Polyak, S.W., Sejer Pedersen, D., Morona, R., Turnidge, J.D., Wallace, J.C., Wilce, M.C.J., et al. (2012). Selective inhibition of biotin protein ligase from Staphylococcus aureus. J. Biol. Chem. *287*, 17823–17832.

Sternicki, L.M., Wegener, K.L., Bruning, J.B., Booker, G.W., and Polyak, S.W. (2017). Mechanisms Governing Precise Protein Biotinylation. Trends Biochemical Sci. *42*, 383–394.

Takayama, K., Wang, C., and Besra, G.S. (2005). Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. Clin. Microbiol. Rev. *18*, 81–101.

Thompson, A.P., Wegener, K.L., Booker, G.W., Polyak, S.W., and Bruning, J.B. (2018a). Precipitant–ligand exchange technique reveals the ADP binding mode in Mycobacterium tuberculosis dethiobiotin synthetase. Acta Cryst D *74*, 965–972.

Thompson, A.P., Salaemae, W., Pederick, J.P., Abell, A.D., Booker, G.W., Bruning, J.B., Wegener, K.L., and Polyak, S.W. (2018b). Mycobacterium tuberculosis dethiobiotin synthetase facilitates nucleoside triphosphate promiscuity through alternate binding modes. ACS Catalysis *8*, 10774–10783.

Tieu, W., Polyak, S.W., Paparella, A.S., Yap, M.Y., Soares da Costa, T.P., Ng, B., Wang, G., Lumb, R., Bell, J.M., Turnidge, J.D., et al. (2015). Improved Synthesis of Biotinol-5′-AMP: Implications for Antibacterial Discovery. ACS Med. Chem. Lett. *6*, 216–220.

Tommasi, R., Brown, D.G., and Walkup, G.K. (2015). ESKAPEing the labyrinth of antibacterial discovery. Nat. Rev. Drug Discov. *14*, 529–542.

Toroney, R., Hull, C.M., Sokoloski, J.E., and Bevilacqua, P.C. (2012). Mechanistic characterization of the 5′-triphosphate-dependent activation of PKR: Lack of 5′-end nucleobase specificity, evidence for a distinct triphosphate binding site, and a critical role for the dsRBD. RNA *18*, 1862–1874.

Warrilow, A.G.S., Hull, C.M., Parker, J.E., Garvey, E.P., Hoekstra, W.J., Moore, W.R., Schotzinger, R.J., Kelly, D.E., and Kelly, S.L. (2014). The Clinical Candidate VT-1161 Is a Highly Potent Inhibitor of Candida albicans CYP51 but Fails To Bind the Human Enzyme. Antimicrob. Agents and Chemother. *58*, 7121–7127.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Cryst D *67*, 235–242.

Woong Park, S., Klotzsche, M., Wilson, D.J., Boshoff, H.I., Eoh, H., Manjunatha, U., Blumenthal, A., Rhee, K., Barry, C.E., Aldrich, C.C., et al. (2011). Evaluating the Sensitivity of Mycobacterium tuberculosis to Biotin Deprivation Using Regulated Gene Expression. PLoS Pathogens *7*, e1002264.

Xu, Y.W., Moréra, S., Janin, J., and Cherfils, J. (1997). AlF3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP. Proc. Natl. Acad. Sci. U.S.A. *94*, 3579–3583.